DOCUMENT RESUME

ED 399 772 FL 024 097

AUTHOR	Local, J. K., Ed.; Warner, A. R., Ed.
TITLE	York Papers in Linguistics, Volume 17.
INSTITUTION	York Univ. (England). Dept. of Language and Linguistic Science.
REPORT NO	ISSN-0307-3238
PUB DATE	Mar 96
NOTE	471p.; For individual articles, see FL 024 098-111.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	York Papers in Linguistics; v17 Mar 1996
EDRS PRICE	MF01/PC19 Plus Postage.
DESCRIPTORS	African Languages; Articulation (Speech); *Autism; Black Dialects; Chinese; *Classroom Communication; Cooperation; Diachronic Linguistics; Disabilities; Echolalia; English; English (Second Language); Finnish; Foreign Countries; French; Grammar; Greek; Group Dynamics; Interpersonal Competence; Italian; *Language Research; Language Rhythm; *Linguistic Theory; Old English; Phonetics; Pronunciation; Second Language Instruction; *Second Languages; Sex Differences; Standard Spoken Usage; Suprasegmentals; Syntax; *Uncommonly Taught Languages
IDENTIFIERS	Gerunds; Kalenjin Languages; *Repairs (Language)

ABSTRACT

These 14 articles on aspects of linguistics include the following: "Economy and Optionality: Interpretations of Subjects in Italian" (David Adger); "Collaborative Repair in EFL Classroom Talk" (Zara Iles); "A Timing Model for Fast French" (Eric Keller, Brigitte Zellner); "Another Travesty of Representation: Phonological Representation and Phonetic Interpretation of ATR Harmony in Kalenjin" (John Local, Ken Lodge); "On Being Echolalic: An Analysis of the Interactional and Phonetic Aspects of an Autistic's Language" (John Local, Tony Wootton); "The Nature of Resonance in English: An Investigation into Lateral Articulations" (David E. Newton); "Prosodies in Finnish" (Richard Ogden); "Old English Verb-Complement Word Order and the Change from OV to VO" (Susan Pintzuk); "Situating 'Que'" (Bernadette Plunkett); "Event Structure and the "Ba" Construction" (Catrin Sian Rhys); "Explanation of Sound Change: How Far Have We Come and Where Are We Now?" (Charles V. J. Russ); "Has It Ever Been 'Perfect'? Uncovering the Grammar of Early Black English" (Sali Tagliamonte); "Voice Source Characteristics of Male and Female Speakers of French" (Rosalind A. M. Temple); and "Notes on Temporal Interpretation and Control in Modern Greek Gerunds" (Georges Tsoulas). (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 399 772

York Papers In Linguistics

17

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Stephen J Harlow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL 024 097

BEST COPY AVAILABLE

York
Papers
In
Linguistics

17

Editors

JK Local
and
AR Warner

ISSN 0307-3238
MARCH 1996

SERIES EDITORS SJ HARLOW AND AR WARNER

DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
UNIVERSITY OF YORK, HESLINGTON, YORK YO1 5DD, ENGLAND

EDITORIAL BOARD

Professor James Hurford (Edinburgh)
Professor John Local
Professor Robert Le Page
Professor Neil Smith (University College London)

David Adger
Connie Cullen
Steve Harlow
John Kelly
Richard Ogden
Susan Pintzuk
Bernadette Plunkett
Charles Russ
Sali Tagliamonte
Ros Temple
Georges Tsoulas
Mahendra Verma
Carol Wallace
Anthony Warner

The Editors are grateful to members of the Editorial Board (and to former members Patrick Griffiths and Joan Russell), who have acted as referees and whose advice has contributed to the quality of the papers published.

Papers submitted to York Papers in Linguistics are each sent to two referees whose anonymous reports must both recommend acceptance for the paper to be published. This issue contains 14 papers; a further ten papers were submitted but were not accepted.

John Local
Anthony Warner

CONTENTS

DAVID ADGER Economy and Optionality: Interpretations of subjects in Italian 1

ZARA ILES Collaborative Repair in EFL Classroom Talk 23

ERIC KELLER AND BRIGITTE ZELLNER A Timing Model for Fast French 53

JOHN LOCAL AND KEN LODGE Another Travesty of Representation: Phonological representation and phonetic interpretation of ATR harmony in Kalenjin 77

JOHN LOCAL AND TONY WOOTTON On Being Echolalic: An analysis of the interactional and phonetic aspects of an autistic's language 119

DAVID E NEWTON The Nature of Resonance in English: An investigation into lateral articulations 167

RICHARD OGDEN Prosodies in Finnish 191

SUSAN PINTZUK Old English Verb-Complement Word Order and the Change from OV to VO 241

BERNADETTE PLUNKETT Situating *Que* 265

CATRIN SIÂN RHYS Event Structure and the *Ba* Construction 299

CHARLES V. J. RUSS Explanation of Sound Change: How far have we come and where are we now? 333

SALI TAGLIAMONTE Has It Ever Been 'Perfect'? Uncovering the Grammar of Early Black English 351

ROSALIND A.M. TEMPLE Voice Source Characteristics of Male and Female Speakers of French 397

GEORGES TSOULAS Notes on Temporal Interpretation and Control in Modern Greek Gerunds 441

EDITORIAL STATEMENT AND STYLE SHEET 471

ECONOMY AND OPTIONALITY: INTERPRETATIONS OF SUBJECTS IN ITALIAN[*]

David Adger

Department of Language and Linguistic Science
University of York

1. Goals

Optional movement is inconsistent with the notion of Economy. Interestingly, optional movement seems to correlate with different interpretations for the resulting structures; when movement is obligatory, on the other hand, the single resulting structure seems to have both of the possible interpretations assigned to the two structures given by optional movement. Why should these facts hold? I provide an answer which is based on the observation that the 'interpretational' differences noticed are actually not semantic at all, but fall within the purview of a separate field of linguistic competence: the ability that human beings have to assign sentences values as to their felicity in discourses. Given this, it follows that there must be an independently specified set of well-formedness conditions deriving well-formed discourses (see, for example work in DRT, especially Kamp and Reyle 1993). I argue that apparent optionality in syntax arises because of a constraint requiring each well-formed discourse to correspond to a collection of corresponding well-formed syntactic structures. Optionality in syntax then becomes essentially a meta-construct, arising out of the interaction between two independent subsystems of

---

[*] Many thanks to the following people for comments on the ideas presented here: Elena Anagnostopoulou; Hagit Borer; Richard Breheny; Itziar Laka; Fabio Pianesi; Manuela Pinto; Bernadette Plunkett; Josep Quer; Tanya Reinhart; Enric Vallduví and Anthony Warner. Many thanks also to Sandra Paoli for help with the data.

York Papers in Linguistics 17 (1996) 1-21
© David Adger

linguistic competence. The apparent interpretational effects are actually effects that arise because native speakers attempt to construct different discourse contexts to satisfy the principles that map between syntax and discourse. The vitiation of these effects when movement is obligatory arises through the interaction of this theory of the interface and the requirement that the syntax be economical. I illustrate this conceptual framework here by taking two narrow domains: subject placement in Italian and the infelicity of anaphoric linkage in discourse across the scope of a quantificational expression.

2. The Problem

Consider the following well-known paradigm from Standard Italian (I shall ignore throughout this paper cases of so called free-inversion where the post verbal subject is not in its theta-position - see Belletti 1988):

(1) Tre leoni hanno sternutito.
three lions have-3p sneeze-pp
'Three lions have sneezed.'

(2) *Hanno sternutito tre leoni.
have-3p sneeze-pp three lions

(3) Tre leoni sono scappati.
three lions be-3p escape-pp-3p
'Three of the lions have escaped.'

(4) Sono scappati tre leoni.
be-3p escape-pp-3p three lions
'Three lions have escaped.'

Assuming some version of the Unaccusative Hypothesis (Perlmutter 1979; Burzio 1985), this paradigm raises an important question for theories of grammar which incorporate some notion of Economy of movement (Chomsky 1989, 1992, 1995): why, if movement is a 'last resort' operation, is (3) a possible syntactic structure? Under the Unaccusative Hypothesis, (4) is essentially the base structure (where the subject is in its theta-position) and there appears to be no motivation for the subject to move to result in (3).

Now consider (3) and (4) more carefully. Belletti (1988) has argued that in (4) there is a definiteness effect which can be seen as long as we make sure that the complement is not free-inverted to a position outside VP. She gives examples with ditransitives:

(5) Ogni studente era finalmente arrivato a lezione.
every student be-3s finally arrived to the lecture
'Every student finally arrived to the lecture.'

(6) *Era finalmente arrivato ogni studente a lezione.
be-3s finally arrived every student to the lecture

Interestingly, as noticed by Pinto (1994), the surface subject position of unaccusatives also shows an interpretative effect. Pinto claims that pre-verbal unaccusative subjects have to be interpreted as being D-linked (Pesetsky 1987); that is they have already been introduced in the discourse. This contrasts with the case of the unergative subject, which has no D-linking constraint imposed upon it.

There are three questions then: why can the subject move? Why does this result in an interpretative difference for the two resulting structures whereby the pre-verbal subject of an unaccusative is D-linked? And why, in the case of unergatives (and transitives) are pre-verbal subjects not necessarily D-linked? (I will ignore the definiteness effect in (6) in this paper, since I think it has an independent explanation.)

3. A Potential Solution

A potential solution to the first problem is suggested by Belletti's (1988) analysis of post-verbal subjects and developments of her ideas by de Hoop (1992) among others. Belletti claimed that the definiteness effect in (5) could be explained by the nature of the type of Case assigned by the unaccusative verb. She terms this Case 'partitive', assumes that its assignment is optional, and correlates it with indefiniteness. De Hoop points out problems with this idea, but essentially develops this line of thought, arguing for different types of Case assignment in the syntax, corresponding with different types of interpretative effect. I shall refer to the hypothesis that the kind of data in (5) and (6) can be dealt with through Case assignment as the Case Determination of Interpretation hypothesis (CDI).

How might the CDI account for the data in (5) and (6)? De Hoop proposes two types of structural Case which she terms 'weak' and 'strong'. For her, these correlate semantically with weak and strong readings of DPs, where a strong reading is essentially a generalised quantifier reading, and a weak one we can take for the moment as existential. Under the CDI we could propose that V-unaccusative assigns weak case to its complement and the auxiliary *essere* assigns strong case to its specifier. This will give us the right interpretative consequences.

What about (1), where the subject can have both interpretations? In this case we could say that the auxiliary *avere* assigns either type of Case to its specifier, which would mean that the subject of an unergative could have either type of reading. Note that if Pinto is right in her semantic characterisation of the readings of subjects in Italian, we can link the notion of D-linked to that of strong Case, and non-D-linked to that of weak Case.

One point of clarification: we cannot actually make the type of Case assigned relate to the auxiliary directly, since the same facts pertain when there is no auxiliary. We must therefore make I bear the Case assigning features, or assume an abstract auxiliary. However, for convenience I will refer to the Case assigning properties of *essere* and *avere* even though actually these properties are instantiated on finite I.

Unfortunately, however, this solution will not generalise effectively to other languages. French is a language which displays similar auxiliary selection facts to Italian and also displays a definiteness effect in impersonal passives:

(7) Il est arrivé trois femmes/ *chaque femme.
it be-3s arrive-pp three women/ *each woman
'There arrived three women/*each woman.'

(8) Trois femmes/ chaque femme sont/est arrivée(s).
three women/ each woman be-3p/be-3s arrive-pp-f(p)
'Three women/Each woman have/has arrived.'

However, French does not appear to display an anti-definiteness effect in (8), which is felicitous in contexts where the subject is non-D-linked. To capture the difference between Italian and French under the CDI one would be forced to jettison the claim that the type of Case was related to the type of auxiliary (or finite inflection) since in (8) we see the equivalent of the *essere* auxiliary in French with either a D-linked or non-D-linked subject.

Furthermore, the CDI seems to miss an important correlation which can be stated in the following intuitive terms: if movement to a position is optional then the two possible structures will have different interpretations; if movement to a position is obligatory, then both interpretations are available for the single structure. This correlation would seem to be essentially functional: you move something to a position to achieve an interpretative effect. In Section 5 of this paper I will develop a formal explanation for the correlation.

In the next two sections I want to present the details of an alternative view to the CDI. I'll argue that the interpretation of preposed subjects of unaccusatives in Italian is not simply that they are D-linked, but rather that such subjects behave as though they are required to be discourse anaphoric (in the sense of Discourse Representation Theory (Heim 1982; Kamp 1981; Kamp and Reyle 1993)). I'll do this by showing that preposed subjects of unaccusatives obey the same constraints as other discourse anaphors such as definites with respect to the scope of adverbial quantifiers (which are discourse anaphor islands). To do this I'll present a version of DRT designed to capture these effects.

I'll then argue that a maximally simple view of Case should be maintained, whereby Case has no interpretative force. It is required to license a DP but not sufficient to determine that DP's surface position. This does away with the notion of optional Case assignment as in Belletti's system. It also paves the way for an explanation of the interpretative correlates of subject placement. The idea is that movement of the subject of an unaccusative to pre-verbal position is an option not because of Case optionality but rather because of conditions regulating the pairing of S-Structures and Discourse Representation Structures. A simple theory of Economy interacts with these conditions to explain the interpretative consequences of optional as opposed to obligatory subject raising.

4. Some Semantics

4.1 A Little DRT

Within Discourse Representation Theory (DRT) indefinites and definites contrast with true quantifiers such as ***every*** in that they are treated as free variables which only become bound during the interpretation procedure. These free variables are termed discourse referents (DRs) and a Discourse Representation Structure (DRS) consists of a universe of DRs and a collection of constraints on those DRs. An example might make this clearer:

(9) a. A man entered. He sat down.
    b. Every man entered. # He sat down.

In (9a) the subject of the first sentence introduces a DR x which is constrained so that the formula **man(x)** must be true of it. Furthermore, the predicate of the sentence, **enter**, must also be true of it. This gives the following representation:

(10)

x
man(x) enter(x)

The pronoun in the second sentence of (9a), being a definite, introduces a further DR y, of which the condition that y sat down must hold:

(10)'

x y
man(x) enter(x) sat-down(y)

Given what I have said so far there does not appear to be any distinction between indefinites and definites. Both introduce DRs and constrain then with formulae. However, in order to capture the fact that the use of a definite pronoun is infelicitous unless there is something for the pronoun to refer back to (I use *refer* here intuitively), Heim (1982) proposes a felicity condition on definites, including pronouns:

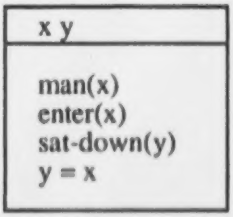
(11) Suppose something is uttered under the reading represented by ϕ (where ϕ is an LF) and the discourse preceding ϕ has resulted in a DRS $\mathcal{K}$, $\mathcal{K}$ contains a set of discourse referents $\mathcal{U}$. Then for every chain C in ϕ it must be the case that:

Familiarity Condition: if C is a definite (including a definite pronoun) then there is a discourse referent x associated with C and $x = y, y \in \mathcal{U}$.

otherwise ϕ is infelicitous with respect to $\mathcal{K}$

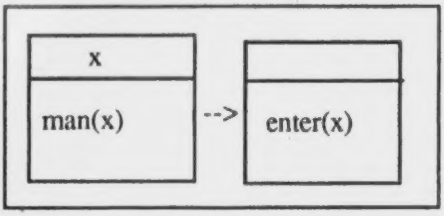
This condition does not hold of indefinites like numerals, *some*, *many*, *several* etc. predicting that indefinites can begin discourses while definites cannot. The Familiarity Condition means that the DRS corresponding to (9a) will actually have to look as follows:

(12)



How then does this theory explain the infelicity of (9b)? The answer is in the DRT structures for quantified sentences (including sentences with adverbial quantifiers - this will become important later on). Kamp (1981) argues that sentences which contain a quantifier give rise to a sub-DRS within the main DRS. The extent of the sub-DRS is defined by the scope of the quantifier. Crucially the DRs in this sub-DRS are not accessible for anaphoric linkage from the main DRS:

(13)



If we were to continue the first sentence of (9b) with the second, then the felicity condition on pronouns (12) will require the DR of the pronoun to be anaphorically linked with a DR in the main DRS. But there is no DR in the main DRS, leading to the correct prediction of infelicity of this sentence with respect to this discourse. I have followed Kamp's early notation for universal quantification here, using an implication sign. In actual fact it will turn out that we need to be specific about the quantificational relation between the two sub-DRSs in structures like (13) - see Kamp and Reyle (1993) for discussion.

Some types of DP always enter their discourse referent in the main DRS though, even if they are in the scope of a quantifier. Examples are proper names and usually definites including demonstratives. So the following is a felicitous discourse:

(14) Every lion in captivity lived in this zoo. We thought it was secure, but they've all escaped now.

Here *it* refers to the zoo, which is possible because demonstratives enter their discourse referents in the main discourse and therefore the felicity condition on *it* can be met. This sentence also illustrates that the plural pronoun *they* seems to be able to pick up a group constructed out of the lions mentioned. The anaphoric properties of plural pronouns lie outside the scope of this paper (but see Kamp and Reyle 1993), but note that *every lion* triggers singular not plural agreement and can be anaphorically picked up by a singular pronoun in its scope, illustrating that something extra is going on with plural pronoun anaphora:

(15) Every lion in captivity wanted its freedom/knew that it needed to be free.

4.2 The Interpretation of Preposed Subjects

Preposed subjects of unaccusatives in Italian[1] appear to behave just like other discourse anaphors, even when they contain a cardinal (indefinite) like *tre* 'three'. Consider the following dialogues:

(16) Questioner: I hear you have lots of cats and dogs staying with you just now. How are they?

Speaker: Tre gatti sono scappati
three cats be-3p escape-pp-3p
'Three cats have escaped.'

#Sono scappati tre gatti.
be-3p escape-pp-3p three cats

---

[1] The judgements here are from Standard Northern Italian.

(17) Questioner: How are you feeling?

Speaker:	Sono preoccupato.	Sono	scappati	tre leoni.
(works in a zoo)	I'm worried.	be-3p	escape-pp-3p	three lions

#Sono preoccupato.	Tre leoni	sono scappati.
I'm worried.	three lions	be-3p escape-pp-3p

With the unaccusative verb it appears that when there is a discourse referent available for ***tre leoni*** 'three lions' then pre-verbal position is the only one allowed. When there is no discourse referent available, then only post-verbal position is felicitous. So far, this squares with Pinto's report and one might imagine an account based on previous mention.

With subjects of unergatives, only pre-verbal position is allowed. We see this below:

(18) Questioner: I hear you have lots of cats and dogs staying with you just now. Have they been up to anything funny?

Speaker: Si,	ieri	tre gatti	hanno sternutito.
yes,	yesterday	three cats	have-3p sneeze-pp

'Yes, yesterday three cats sneezed.'

(19) Questioner: Have you seen anything funny lately?

Speaker: Si,	ieri	tre gatti	hanno sternutito	lungo la strada.
yes	yesterday	three cats	have-3p sneeze-pp	along the street

'Yes, yesterday I saw three cats sneeze on the street.'

Note that in contrast to (17) the pre-verbal position is fine whether there is an available discourse referent or not. Again this seems to follow Pinto's claim that D-linking is irrelevant for unergative subjects.

However, there is an argument that DRT style accessibility is actually what's at stake here, rather than just previous mention in the discourse. Consider the following two discourses:

(20) a. Ogni volta che le pop-stars e i divi del cinema che vivono al numero 27 ritornano a casa, mi emoziano.
'Every time the pop-stars and film stars that live at number 27 come home, I get excited.'

b. Ieri, tre pop-stars sono arrivate.
yesterday, three pop-stars be-3p arrive-3pf
'Yesterday, three of the pop-stars came back.'

b'. Ieri, sono arrivate tre pop-stars.
yesterday be-3p arrive-3pf three pop-stars
'Yesterday, three pop-stars arrived.'
(must be different pop-stars from those living at no. 27)

(21) a. Ogni volta che delle pop-stars venguno nella mia strada, mi emoziano.
'Every time pop-stars come to my street, I get excited.'

b. #Ieri, tre pop-stars sono arrivate.
yesterday, three pop-stars be-3p arrive-3pf
'Yesterday, three of the pop-stars came back.'

b'. Ieri, sono arrivate tre pop-stars.
yesterday, be-3p arrive-3pf three pop-stars
'Yesterday, three pop-stars arrived.'

In both of these sentences we have an adverbial quantifier which will give rise to sub-DRSs in DRT. This predicts that discourse referents that are inside the scope of the quantifier are not accessible to those outside. In (20a), however, we have a definite, which is entered in the topmost discourse and a pre-verbal subject in (20b) is well-formed. A post-verbal subject (20b') is also well formed, on the condition that the pop-stars referred to are not the ones previously introduced (the familiar definiteness effect). In (21a), the discourse referent of pop-stars is introduced by an indefinite, it will therefore be interpreted within the scope of the quantificational adverb predicting that it is not accessible for anaphoric reference. Given this, to predict the infelicity of (21b), we simply need to say that whatever is in the specifier of IP falls under the Familiarity Condition given above in (11) and repeated here.

(22) Suppose something is uttered under the reading represented by ϕ and the discourse preceding ϕ has resulted in a discourse structure $\mathcal{K}$. $\mathcal{K}$ contains a set of discourse referents $\mathcal{U}$. Then for every chain C in ϕ it must be the case that:

Familiarity Condition: if C is definite or in Spec, IP[2] then there is a discourse referent x associated with C and $x = y$, $y \in \mathcal{U}$.

otherwise ϕ is infelicitous with respect to $\mathcal{K}$

The point about (21) is that (21a) creates a sub-discourse $\mathcal{K}$ the discourse referents of which are not accessible except within $\mathcal{K}$. (21b) however, is outside $\mathcal{K}$, but contains an element in Spec, IP. There is no discourse referent in $\mathcal{U}$ which the discourse referent of *pop-stars* can be equated with. (21b) is therefore infelicitous with respect to (21a).

4.3 Mapping between Syntax and DRS

Note that the condition $x=y$ is essentially non-linguistic. Definites behave in exactly the same way with respect to anaphora and deixis (Kartunnen 1976) so if we wish to capture this fact we need to assume that such a condition can be entered into the DRS non-linguistically, by an act of ostension, or something similar. This point is crucial, in that it means that there must be independent well-formedness conditions on the construction of DRSs.

---

[2] I have formulated the Familiarity Condition here using the notion Spec IP. This is only for reasons of exposition, and readers will recognise that there is an issue as to exactly what kind of syntactic description should go in here so as to capture the widest variety of data. In Adger 1994 I developed the notion of Agr-Chain, which is a chain with a link in Spec AgrP and argued that by using this notion in the Familiarity Condition one could unify the interpretative effects that arise with subject placement, scrambling, clitic-doubling, wh-agreement and case.

The picture of the grammar built up here claims then that there is some set of well-formedness conditions on DRSs, and an independent set of well-formedness conditions on terminal syntactic structures (TSS), where by terminal syntactic structures I mean structures which satisfy all of the constraints of the syntax. TSS then is LF or SS depending on which you take to be the input to interpretation. Felicity conditions like the Familiarity Condition are essentially relations between DRSs and TSSs. Further mapping principles link other aspects of TS structure to aspects of DRS structure (possibly also stipulated in terms of chains). A minimal theory would relate head-chains to predicates in the DRS, and XP chains to DRs.

Are all of these mapping principles of the form F(TSS)=DRS? Are there any constraints the other way round? That is, are there mapping principles which are of the form F(DRS)=TSS? I would like to suggest that there is at least one and that it is this principle rather than Case which motivates movement of a subject of an unaccusative to Spec IP position. This principle essentially claims that the non-linguistically introduced information in a DRS must also be able to be linguistically introduced.

Assume that the (infinite set) of DRSs given by the DRS well-formedness conditions is $\mathcal{P}$, and the set of TSSs given by the syntax is $\mathcal{L}$, then:

(23) **Effability:** For every member p of $\mathcal{P}$ there is a corresponding member ℓ of $\mathcal{L}$

where ℓ corresponds to p iff for every felicity condition F, $F(\ell)=p$.[3]

5. Some Syntax

5.1 Movement and Economy

Chomsky (1991, 1992, 1995) has recently proposed that a number of grammatical principles might be reduced to principles governing the complexity of derivations and representations, where complexity is to be theoretically pinned down. For example, the principle of 'least-effort' requires that a derivation must be as 'short' as possible deriving the effects of the ECP under a relativised minimality view of the latter (Rizzi 1990). A further principle of Economy prohibits operations which are not needed to enable the derivation to successfully converge. For my purposes, it is sufficient to propose a rather general theory of Economy, of the following sort:

---

[3] Fabio Pianesi has pointed out to me that this definition as it stands will not halt. This problem can of course be solved trivially by requiring a single pass in whatever algorithm is used to implement it.

(24) **Economy:**
Minimise computational operations

Computational operations are copying, insertion and deletion as in the earliest versions of transformational grammar (Chomsky 1955). I will assume that movement consists of (one or more) copying operations, followed by a deletion operation, as argued in Chomsky (1992). Note that deletion may take place at TSS to satisfy the requirements of Full Interpretation (as discussed in Chomsky 1992 for reconstruction effects) or at PF (perhaps for cases of ellipsis, etc.). Deletion is of course subject to recoverability of content.

This theory of Economy should be construed globally, in the sense of Reinhart (1994) and Adger (1995). That is, a derivation leading to a particular TSS will be deemed to be more expensive than another derivation leading to the same structure if the former consists of more computational operations. It is in this sense that computational operations should be minimised.

5.2 Capturing the correlations

Let us return to our original paradigm (repeated here):

(25)	Tre	leoni	hanno	sternutito.
	three	lions	have-3p	sneeze-pp
	'Three lions have sneezed.'			

(26)	*Hanno	sternutito	tre	leoni.
	have-3p	sneeze-pp	three	lions

(27) Tre leoni sono scappati.
three lions be-3p escape-pp-3p
'Three of the lions have escaped.'

(28) Sono scappati tre leoni.
be-3p escape-pp-3p three lions
'Three lions have escaped.'

Ideally we would like to capture this with a minimal theory of Case, something like the following:

(29)
- V assigns Case to its complement, and not to its specifier.
- I assigns Case to its specifier.

This theory predicts that an unaccusative subject gets Case in its theta-position (complement of V position in (28)), and an unergative subject must move to Spec IP ((25) - because it cannot get Case in Spec VP, assuming that is its theta-position (Koopman and Sportiche 1991)). Ignoring Economy, it also predicts that a Spec IP subject of an unaccusative verb is well-formed ((27) - since it can receive Case there from I), and that a post-verbal subject of an unergative is bad (since it doesn't get Case - (26)). However, given Economy, why will an unaccusative subject ever raise to Spec IP if it can get Case in its theta position?

The answer Belletti (1988) proposes is that the Case assigned by unaccusatives is always optional. When the option is not taken to assign Case, then the subject must raise to Spec IP to get Case there.

There is an alternative solution which does not involve complicating Case theory in this way. An unaccusative subject will raise if there is some further well-formedness principle that it must obey. Now, note that if (27) were ill-formed there would be no TSS corresponding to the DRS where the DR of the subject is a discourse anaphor. This is in violation of Effability, which requires that for each DRS there be a corresponding TSS. Effability then requires that (27) be a possible TSS of Italian (note that to make this story go through, we have to assume that TSS is S-Structure for Italian. I suspect that it's S-Structure for all languages).

To see how this works in more detail consider the schematic structures of (27) and (28):

(30) a. escape three lions (nothing in Spec IP)
b. three lions escape (*three lions* in Spec IP)

The question is why (30b) is well-formed. (30a) corresponds to a DRS with a single plural discourse referent (say x) and three conditions on that discourse referent: **lion(x)**, **three(x)** and **escape(x)**. This DRS is given independently by the DRS well-formedness conditions.

(30b) is a possible TSS because Effability requires there to be a TSS corresponding to a DRS where the escaping lions are anaphoric to some previously established lions. This will only be true if there is a TSS of which the Familiarity Condition holds for the three lions. This in turn will only be true if the DP *three lions* is definite or is in Spec IP. But surely this predicts that we can simply make the DP definite, rather than move it to Spec IP.

This conclusion certainly follows given what we have said so far. However, the felicity conditions on definites and those on Spec IP elements appears to be different. Crucially, it is possible to accommodate (that is to use a definite which hasn't itself been introduced in the discourse but is inferable from the discourse) from a definite in post-verbal position but not from pre-verbal position (see also Anagnostopoulou 1994 who first pointed out similar facts concerning clitic doubling in Modern Greek, and see Delfitto 1994 for scrambling of objects in Dutch):

(31) Ieri ho visto un film su Fellini,
'Yesterday I saw a film about Fellini,'

a. e oggi e arrivato il regista a casa mia.
and today be-3s arrive-3s the director to my house
'and today the director (of the film) arrived at my house.'

b. e oggi il regista e arrivato a casa mia.
and today the director be-3s arrive-3s to my house
'and today the director (Fellini) arrived at my house.'

Given this we need to tease apart the Familiarity Condition into two sections, where one part regulates Spec IP elements and the other regulates definites.

Then Effability forces the syntax to generate (27), even though (28) is well-formed.

The next question is why (27) is only felicitous with a discourse anaphoric reading for its subject, while (25) is felicitous with a discourse anaphoric reading or not. The answer to this question is the interaction of Economy with Effability.

Note that there are actually two chains that result from raising an unaccusative subject into Spec IP (30b) under the copy-and-delete view of movement outlined above, depending upon which copy is deleted. I will for the moment stipulate that (30b) itself is not a TSS and that either the link in Spec IP or the link in Compl VP must be deleted. This requirement is probably derivable from the different Mapping Conditions on VP internal and VP external objects, but I shall not go into that here (see Adger 1994, 1995; Diesing 1992). If we delete the copy in complement of V position we have an element in Spec IP, while if we delete the copy that is in Spec, IP position we obviously have nothing in Spec IP:

(32) a. ~~a lion~~ escape a lion
b. a lion escape ~~a lion~~

This would appear to predict that a preposed subject of an unaccusative would have two readings, since there appear to be two TSSs for this sentence, contrary to fact.

However, note that the derivation of (32a), the variant where three lions is not discourse anaphoric involves two computational operations: Copy α, followed by Delete α. Note also that the result of this two-step derivation is exactly the result of not raising the subject in the first place. Given the theory of Economy discussed above, we predict that (32a) is not actually a TSS for (30b). So a raised subject of an unaccusative verb does not have a non-discourse anaphoric reading, because the derivation that would give rise to that reading is blocked by the existence of an alternative structure which involves less computational steps.

In contrast consider the schematic form of an unergative:

(33) a. three lions sneeze
b. * sneeze three lions

The simple Case theory outlined in (29) rules out (33b). Given the discussion above, however, we still have two putative TSSs for (33a):

(34) a. ~~three lions~~ sneeze three lions (nothing in Spec IP)
b. three lions sneeze ~~three lions~~ (*three lions* in Spec IP)

Note that there is no competing derivation in this case for (34a) since (33b) is ruled out anyway. This predicts that the subject of an unergative verb will have both readings, as it does.

5.3 A potential problem

The system outlined so far predicts that when movement to a position is optional then a structure involving the moved element will have a different interpretation from the structure involving the *in-situ* element. Specifically, with subject placement, it predicts that when a VP internal position for the subject is available, as well as Spec IP, then Spec IP subjects will be discourse anaphoric. An empirical problem for this prediction appears to arise in Catalan. In Catalan the canonical subject position for all verbs appears to be VP-internal (Vallduví 1993). An unergative verb like *trucar*, 'phone', allows a post-verbal subject and is felicitous in discourses where the subject is discourse anaphoric or not (again controlling for right dislocation):

(35) a. Deuran trucar alguns convidats, oi?
must-3p call some guests, right
'Some (of the) guests will probably call, right?'

Note that there is no definiteness effect here, even though the subject is VP internal. This contrasts with Italian, suggesting that the definiteness effect in Italian relates to a null expletive in subject position, which is not present in Catalan. The subject can also be preposed:

(35) b. Alguns convidats deuran trucar, oi?
some guests must-3p call, right
'Some (of the) guests will probably call, right?'

Unfortunately, there appears to be no interpretational difference here, contrary to the predictions of the theory.

However, there is an independent explanation for this effect. Catalan actually seems to have two subject positions: Spec IP, and an IP adjoined position. Vallduví (1992) has argued that Spec IP in Catalan is reserved for quantificational elements on a weak reading (that is in our terms non-discourse anaphoric). Vallduví argues that referential elements are barred from this position. The IP adjoined position, on the other hand, corresponds to the subject position in Italian and must be interpreted as discourse anaphoric.

6. Conclusion

This paper has argued that subject placement in Italian is not entirely determined by Case, but rather that it is also partly determined by interpretational considerations. The crucial step in the argument is that there are independent well-formedness conditions on discourse structures and that the apparent interpretational effects on preposed subjects of unaccusatives in Italian are actually effects that derive from judgements of felicity in discourse. The apparent optionality of syntactic movement is in fact conditioned by an interface constraint that requires each well-formed DRS to have a set of corresponding terminal syntactic structures. These considerations interact with a notion of global Economy to derive the correlation between subject placement, optionality and interpretation.

This conclusion actually reinforces the autonomy of syntax rather than threatens it. It removes any features from the syntax which have purely interpretational motivation and leaves a simple theory of argument licensing which is purely structural.

REFERENCES

Adger, D. (1994) *Functional heads and interpretation*, Ph.D., University of Edinburgh.

Adger, D. (1995) Meaning, Movement and Economy. In R. Aranovich, W. Byrne, S, Preuss, and M. Senturia (eds.), *Proceedings of WCCFL 13*. Stanford: CSLI.

Anagnostopoulou, E. (1994) *On the representation of clitic doubling in Modern Greek*, ms., University of Tilburg.

Belletti , A. (1988) The Case of Unaccusatives *LI* 19.1-34.

Burzio, L. (1985) *Italian Syntax*. Dordrecht: Reidel.

Chomsky, N. (1955) *The Logical Structure of Linguistic Theory*. New York: Plenum.

Chomsky, N. (1989) Some notes on economy of derivation and representation *MIT Working Papers in Linguistics* 10.43-74.

Chomsky, N. (1992) A minimalist program for linguistic theory. In *Occasional Papers in Linguistics, 1*. MIT.

Chomsky, N. (1995) Bare Phrase Structure. In G. Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell. 385-439.

Delfitto, D (1994) Beyond specificity: Proposals on cliticization and scrambling. Handout from talk given at the Third Plenary ESF Conference on Language Typology, le Bischenberg.

Diesing, M. (1992) *Indefinites*. Cambridge, Mass: MIT Press.

Heim, I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D., University of Massachusetts, Amherst.

Hoop, H. d. (1992) *Case Configuration and Noun Phrase Interpretation*, Ph.D., University of Groningen.

Kamp, H. (1981) A theory of truth and semantic representation. In J. A. S. Groenendijk, T. M. V. Janssen and M. B. J Stokhof (eds.) *Formal Methods in the Study of Language*. Amsterdam: Mathematical Center Tracts.

Kamp, D. and U. Reyle (1993) *From Discourse to Logic*. Dordrecht: Kluwer.

Karttunnen, L. (1976) Discourse referents. In J. McCawley (ed.) *Syntax and Semantics 7*. New York: Academic Press. 363-385.

Pinto, M. (1994) ***Subjects in Italian: distribution and interpretation***, ms. University of Utrecht.

Perlmutter, D. (1979) Impersonal passives and the unaccusative hypothesis *BLS* IV.157-189, University of California.

Pesetsky, D. (1987) Wh-in-situ: movement and unselective binding. In E. J. Reuland and A. G. B. ter Meulen (eds.) *The Representation of Indefiniteness*. Cambridge, Mass: MIT Press.

Reinhart, T. (1994) ***From syntactic encoding to interface strategies***, ms, OTS, Utrecht.

Rizzi, L. (1990) ***Relativized Minimality***. Cambridge, Mass: MIT Press.

Vallduví, E (1992) A preverbal landing site for quantificational operators. In ***Catalan Working papers in Linguistics***, Barcelona.

Vallduví, E. (1993) ***The informational component***, Ph.D., University of Pennsylvania.

COLLABORATIVE REPAIR IN EFL CLASSROOM TALK

Zara Iles

Department of Language and Linguistic Science
University of York

1. Preface

This paper explores some of the benefits to be gained by adopting a conversation analysis (CA) perspective in an examination of 'English as a foreign language' (EFL) classroom talk. The EFL classroom is a context in which there is a heightened potentiality of problematic talk, e.g. errors, misunderstandings and non-communication. The need for REPAIR (Schegloff et al 1977) is therefore situationally endemic. In everyday talk, between participants who hold mutual assumptions of common ground and shared knowledge, repair has been shown to be an activity which is executed quickly as repair trajectories can necessitate certain interactional investments. EFL teachers and learners are differentially capable of dealing with and resolving trouble-at-talk situations because of the unequal knowledge distribution that exists between them. Some of the ways in which talk created by EFL participants is collaboratively built in order to address this particular state of affairs are discussed in this paper.

It is seen that differences in the agenda of the lesson at hand, e.g. involving a focus on language form or creation of conversation, are reflected in the interactional structure. Forms of correction are shown to impose different costs on the interaction, lesson agenda and for second language learners. Teachers are seen to be orienting to the status of other-correction as the least preferred repair trajectory (Schegloff et al. 1977), by a) pursuing repair initiation, b) withholding correction and c) adopting various camouflages which serve to downgrade the dispreferred activity of other-correction.

York Papers in Linguistics 17 (1996) 23-51
© Zara Iles

1.1 Introduction

This paper arises as part of a larger investigation which examines the ways, and the extent to which, matters pertaining to the development of language competencies are worked on by EFL teachers and learners in their talk. One such matter concerns errors and their treatments, one of the major businesses in which EFL classroom participants routinely engage. In spite of the fact that correction is an activity which is customary in the EFL context, "so little is known about the nature of correction as it occurs in the classroom and its effect on the learning process" (Pica 1994:70). Error and error correction are important in the characterisation of the nature of talk generated between EFL teachers and learners, and as such, a valid and accurate account of this aspect of EFL talk is of primary concern to second language acquisition (SLA) research.

In SLA research deciding on a definition of 'error' and identifying errors has proved problematic. An error is typically, and restrictively, defined as "the production of a linguistic form which deviates from the correct form" (Allwright and Bailey 1991:84); the correct form being that of the native-speaker 'norm'. Lennon (1991) concludes that:

> 'no universally applicable definition can be formulated, and what is to be counted as an error will vary according to situation, reference group, interlocutor, mode, style, production pressures' (Lennon, 1991:331)

A CA approach avoids such categorisation and analyses which result from an investigator's own intuitive understanding of what is happening in an instance of talk. It gives rise to an analysis which is based on observation of the orientations of the participants themselves in creating, and making sense of, their talk. The CA concept of repair allows for a broader perspective of error and correction than what is currently prevalent in SLA research. Repair is the structural and organisational mechanism in conversation that allows speakers to deal with troubles in speaking, hearing or understanding ongoing talk (Schegloff et al 1977). The term thus refers to a wider range of events than simply that of correction, which is just one possible realisation of repair. Repair organisation offers all-inclusive and thus potentially more useful notions of the terms 'error' and 'correction', referring to all instances of problematic talk and the trajectories which are involved in its treatment. Construed in this fashion, errors can thus be seen as being more than the production of a deviant form by the learner, and hence specifically the learner's problem; errors and their repair constitute an interactional problem which EFL participants must jointly overcome, and which involves them in the regeneration of their talk after trouble or breakdown.

Repair entails making some aspect of language the focus of the talk to one degree or other, i.e. correction becomes the explicit activity of the talk or is a 'by-the-way-occurrence' and is dealt with swiftly (Jefferson 1987). Repair sequences are environments in which the identities of the participants as 'teacher' and 'learner' are made interactionally relevant and so manifested in the details of the talk. Repair trajectories are also environments within which knowledge (possibly new knowledge) about the target language is made available for the learner by the teacher. Language is demonstrated, experienced and worked on by both teacher and learner in repair trajectories. As will be shown in this paper, the structure and design of repair trajectories means that the extent of this 'working on talk' is negotiated. A detailed examination of these features of EFL interaction is therefore likely to yield important insights into the nature of second language (L2) development and the nature of its relationship to interaction.

This paper concentrates primarily on other-correction, the least preferred trajectory in repair organisation in everyday talk. Schegloff et al (1977) demonstrate that mundane conversation is 'structurally skewed' so that self-repair opportunities, where the originator of the trouble repairs his/her own talk, dominate over other-repair opportunities, where a co-participant actions the repair. Other-corrections are the forms of repair which Schegloff et al suggest operate as:

> a device for dealing with those who are still learning or being taught to operate with a system which requires, for its routine operation, that they be adequate self-monitors as a condition of competence. It is, in this sense, only a

> transitional usage, whose supersession by self-correction is continuously awaited. (1977:381)

The paper reveals how the recurrent features of repair observed in everyday conversation between native speakers, are employed in a 'specialised' way by participants in the context of the EFL classroom. It further reveals how the forms of repair employed by the EFL teachers, which orient to the maximisation or minimisation of explicit error correction, reflect the nature and the agenda (local and global) of the teaching activity. It also shows that the extent to which error correction becomes the overt business of the talk, or not, can, potentially, be controlled by both teacher and learner. For example, the design of teacher other-correction may serve to downgrade the activity in order to interrupt the ongoing talk as minimally as possible. Various camouflaging features drawn from observing teacher other-correction are highlighted in the extract analyses in section 4. The interaction in which EFL participants are engaged can be designed to either give priority to the business of 'creating conversation', or, the correction of talk and conscious analysis of the target language.

The account given in this paper is developed from observations made by Jefferson (1987) concerning explicit and embedded other-repair and subsequent projected accountings in normal everyday conversation. Examination and discussion of these repair trajectories is presented in Section 2. Instances of these two forms of other-correction from naturally-occurring EFL classroom data are described and discussed in Section 4. It is demonstrated that repair strategies adopted by EFL interactants can synchronously, a) attend to the nature, or expedite the achievement, of different goals to be attained in EFL lessons, and b) be sensitive to the linguistic, cognitive and interactional loads placed on 'less than fully competent' participants.

2. Exposed and Embedded Correction

Jefferson (1987) identifies and describes two forms of other-correction observable in everyday talk which have different interactional consequences; exposed and embedded correction. Jefferson demonstrates that correction by other-speaker is an activity which can either be a)

accomplished explicitly, where the correction becomes the interactional business, or, b) accomplished without it emerging to the conversational surface. Exposed correction has an interactional cost as the ongoing talk is interrupted and correction becomes the concern of the talk. It is demonstrated that with exposed forms of correction:
'correcting can be a matter of, not merely putting things to right ... but of specifically addressing lapses in competence and/or conduct' (Jefferson 1987:88).

After exposed correction, giving an account of error is potentially relevant. Exposed correction may therefore be a means of specifically bringing a participant to account for their errors. On the other hand, embedded other-correction is a way of handling problematic talk without invoking the apparatus of repair, i.e. initiation attempts, repair markers, hesitation, lengthy trajectories and so on, which lead to the successful, or otherwise, treatment of the repairable. Embedded correction does not project accountings and does not discontinue the ongoing talk. Correction does not become the interactional business and therefore demands less interactional investment, less time, and talk stays on topic. The following examples A-D from Jefferson's 1987 paper illustrate these two types of other-correction forms:

(Example A): Other-correction in next-turn with no overt markers (in line 1) and a minimal receipt of correction (in 2). The repairable item is picked out by Norm and an isolated repair, without surrounding syntactic context or explicit repair markers, is performed. The repair is imitated by Norm, marked with stress and acknowledged with an explicit receipt; 'Right'. The correction does not become topicalised, is executed quickly and so the talk is minimally interrupted. The redoing and completion of the repairing is signalled with a minimal 'M-hm' receipt from Norm who actioned the repair.

```
    Larry: They're going to drive back
                     Wednesday
1   Norm:            Tomorrow.
2   Larry: Tomorrow. Righ[t.
3   Norm                     [M-hm,
    Larry: They're working half day.
```

(Example B): Other-correction in next-turn with no overt markers (in 1) and an embedded receipt of repair (in line 2). No account of the error is given by Milly and she continues on topic. In next-turn after the trouble-source turn an other-correction is actioned by Jean. The repairable is isolated, redone without interval or explicit repair markers. The initial consonant is stressed and this is imitated by Milly in her subsequent redoing. Unlike in example A there are no acknowledgement markers of the repair activity from either speakers. The correction proceeds as a by the way occurrence and does not become the explicit focus of the talk.

Milly: ...and then they said something about Kruschev has leukemia so I thought oh it's all a big put on.
1 Jean: Breshnev.
2 Milly: Breshnev has leukemia. So I didn't know what to think.

(Example C): An example of other-correction in next-turn with no overt markers (in 1) and an explicit receipt of correction (from 2 onwards). Jo actions the repair in line 1 without delay and without explicit repair markers. The repair is redone by Pat and she then maintains the repair as the focus of the talk by doing an accounting. Correction becomes the concern of the talk and there is some delay to the topic. The repair activity is made the source of a joke, which orients to the status of other-correction as a dispreferred activity and is a face-saving device.

Pat: ...the Black Muslims are certainly more provocative than the Black Muslims ever were.
1 Jo: The Black Panthers.
2 Pat : The Black Panthers. What'd I
Jo: You said the Black Muslims twice.
Pat: Did I really?
Jo: Yes you di:d but that's alright I forgive you.

In examples A, B and C, the repairable is isolated in the correction turn i.e. there is no surrounding syntactic context. There are no explicit repair markers and the repair is imitated immediately by the originator of the trouble source in the following turn. The repair is executed quickly and there is little interruption to the ongoing talk. The examples also exhibit various behaviours by which participants acknowledge that repair is being accomplished, e.g., intonational highlighting of the repair elements and various minimal receipts. These same features are found in the repair sequences from EFL lessons discussed below in section 4. These sequences were taken from lessons or points in lessons where making correction the focus of talk is not the primary agenda. Explicitly packaged, exposed correction would interrupt the topic and potentially take over as the focus of the talk. The repair structure of examples A and B ensures that a) talk is repaired b) a redoing by the originator of the trouble-source is projected and accomplished, hence this can be regarded as an orientation to self-repair preference in the last resort, and c) the cost of repair activity to the interaction is limited.

The two forms of other-correction highlighted in the examples above do not correspond to two symmetrically distinct modes of correction. Correction may be explicitly actioned by one participant, but be accepted in an embedded form by the co-participant, thus ignoring the potentially projected accounting for error. Likewise, a correction may take an embedded form but be brought to the conversational surface by an explicit receipt. This phenomenon is illustrated in the following example in which participants deal with racist language.

(Example D): Other-correction in overlap (in 1) with explicit repair markers and embedded receipt of correction (in 2).

Jim: Like yesterday there was a track meet at Central.Ree:se was there. Isn't that a reform schoo:l,

(0.4)

Jim: Ree:se?

(.)

Roger: Ye:s.

Ken: [Yeah.
Jim: [Buncha niggers and everything?
Ken: Yeah.
(0.3)
Jim: He went right down on that fie:ld
like a nigger and all the guys
(mean) all these niggers are a:ll
[up there in-]
1 Roger: [You mean Ne]gro: don't you.
(.)
Jim: Well and [they're all-h-u]=
Ken: [And Ji:g,]
Jim: =-They['re they're A:LL up in the
Ken: [hunh stands you know all
(.)
Jim: Th:ese guys (are) completely
2 radical.I think I think Negroes are
cool gu:ys you kno:w,
Ken: Some of them yeah.

In the example above, Roger's exposed correction, in line 1, projects a potential accounting. But the repair is receipted in an embedded form by Jim later in the talk, in 2, thus avoiding having to give an account for his repairable. In this way, Jefferson argues, the activity of correction is shown to be a collaborative enterprise as it is through the participants': 'collaborative, step-by-step construction that correction will be an interactional business in its own right, with attendant activities addressing issues of competence and/or conduct or that correction will occur in such a way as to provide no room for accounting.' (Jefferson 1987:99)

In the EFL classroom context the capacity for this co-operative enterprise is potentially constrained. Second language learners may not be aware of the need for repair, let alone be in a position to action repair for themselves. Consequently, forms of correction may prove to have further costs for L2 teachers and learners. Exposed correction (initiation and treatment) and its accompanying activities can require the learner to focus explicitly and consciously on the form of the language s/he is trying to learn. The learner may not be in a position to be able to meet these projected demands. On the other hand embedded forms of correction empowers the EFL teacher to attend to the repair of trouble-sources, but does not oblige an explicit of consciously motivated focus on language form. The L2 may, if in possession of necessary knowledge, accept the correction in an exposed receipt and even make the correction the focus of the talk him/herself. The continuum of repair and control of preference is negotiated as talk unfolds. For example, where the learner displays no awareness of error or inability to action self-repair in their talk EFL teachers may action other-correction in either an exposed or embedded form. (The employment of these structures is shown in section 4 to be indexical of the pedagogical agenda of the lesson). What is projected as a relevant next is therefore controlled, to some extent or other, by teacher and learner.

The extracts that follow reveal how types of correction are indexical of the agenda of the lesson and learner competence. They also show how various features in the talk of EFL teachers downgrade the activity of other-correction, the least preferred trajectory in the organisation of repair in mundane conversation.

3. Data

The extracts discussed below were selected from a corpus which includes data from audio-taped lessons from 10 native-speaker EFL teachers and 12 learners (of various nationalities). The lessons which were either described as 'conversation classes' or 'business English' took place in language units/schools in York and London. Teachers and learners were not informed of the express purpose of the study and the researcher was not present during the recordings. Factors such as age or sex of the participants were not a pre-consideration of the study reported in this paper and were therefore not controlled for the purposes of the study. Schegloff (1992) states that categorising speakers is only relevant when interactants themselves orient to such distinctions and can be found in the details of the talk. Such information would therefore only be brought to light after analysis of the data. However, some information about the learners and the language schools, where known, is given, and a brief description of the nature of each lesson.

ZLI:SFM:C1
A 'conversation class' at the University of York involving sixteen learners of various nationalities. This class which ran throughout a nine week term was targeted at overseas students and their partners who sought conversation practice. In this lesson the learners, in pairs, have been completing a gap-fill grammar exercise from a textbook. The exercise involves choosing the correct phrasal verb from a range of six possibilities. Extract 1 is taken from the point in the lesson where the whole class is collectively going through answers and correcting mistakes.

ZLI:SFM:GB1
A one-to-one 'conversation class' at the University of York involving a female Turkish native-speaker. The student was enrolled on a course of general English lessons prior to taking pre-sessional EAP courses before the beginning of the academic year. In this lesson the teacher and learner are involved in a discussion of images of Turkey after independently watching a television programme during the week prior to the class and discussing newspaper articles.

ZLI:SFM:P1
A one-to-one 'business English class' at a private language school in the city of York involving a Portuguese native-speaker. At the beginning of this lesson the teacher presented and explained various target sentences for 'comparing and contrasting' and 'giving opinions'. The teacher and learner discuss various statements given in their textbook, the learner's task being to give his opinion about what the statements suggests and to try to employ some of the target language previously given. Examples of statements are "business failure is due to bad management" and "high levels of unemployment will continue for decades".

ZLI:DC:G1
A one-to-one lesson at a private language school in London involving a German native speaker. The teacher and learner are discussing various topics, e.g., theatre, books, television. Some correction is actioned during the course of the conversation as errors occur, but 5 minutes is given over to highlighting errors and working through them at the end of the lesson.

ZLI:A:L1
A one-to-one 'Business English' lesson at a private language school in York. The learner is a French native speaker who is on a one-week course. The lesson was recorded on the last day of the learner's course and the activity in the lesson involves correcting sentences prepared previously for homework and reviewing new language.

4. Analysis of Data Extracts[1]

Extract 1: ZLI:SFM:C1

```
1   T: Horiyo can you read out what you've got
2      for that please.(*) the whole sentence
3   H: Mm hm the local supermarket has got up
4      the pri:ces again
5              (*)
6   T: .HHHh now it's. ([*)       ] the verb
7   L:                   [unintell)]

8   T: is- yes something up yes
9              (*)
10  T: Now what do we sa- [(*)       ] not the
11  L:                    [(unintell)]

12  T: correct verb ([*)        ] no Forget get
```

---

[1] The notation employed in this paper is taken from Atkinson and Heritage (1984). Square brackets indicate the onset and offset of overlapping talk; untimed pauses are marked as (*).

14 L1: G-et
15 L2: -get
16 T: No Forget get p'
17 L: ((unintell))
18 T: What?
19 L Put
20 T: We:ll done good

This first extract is from a lesson where language form and revealing linguistic knowledge is the explicit focus of the talk. Repair is therefore integral to the agenda of the lesson. The teacher nominates a particular learner, H, to make a public display of his competence. The learner provides an incorrect answer. The following delay. (line 5). and in-breath, dispreference markers at the start of the teacher's turn in line 6 signals inability to provide affiliative talk and that further work is needed. Another learner offers a possible answer (unintelligible to the observer). The teacher's turns from line 6 onwards involve repeated other-repair initiation and a marked withholding of other-correction. T highlights where the learners' attempts have been correct, "yes something up yes", in line 8. This initiation does not lead to successful learner repair. No possibles are offered by the learners. The teacher still does not action a correction at this point, but pursues initiation and providing clues. T proceeds to explicitly state that the learner's have chosen an incorrect verb. Further incorrect attempts are forthcoming from the class. In line 16, the teacher gives a further clue "p" to locate the correct verb - 'put' is the only verb in their list beginning with 'p'. The teacher's explicit initiation succeeds in enabling the learners to action the repair for themselves. Although the teacher has avoided unmodulated other-correction, the various steps in the repair initiation has demanded investment in the talk and of the learners' level of linguistic knowledge. The withholding of other-correction and involved repair trajectories to be found in this lesson echo observations made by McHoul concerning repair organisation in subject classroom talk. A regular pattern observed in McHoul's data was for the teacher to reformulate questions as further repair initiation and to provide clues to assist learner self-repair. McHoul concludes that "contrary to what may be a popular image of the classroom, teachers tend to show students *where* their talk is in need of correction, not how corrections should be made" (1990:376). And in showing where, teachers indicate, of course, candidate 'whats'

Extracts 1, 2, 3 and 4 are taken from a lesson where creating conversation is the global pedagogic focus of the talk. The repair in the next extract involves the treatment of a single lexical item by the teacher after no display of error awareness by the learner.

Extract 2: ZLI:SFM:GB1

```
1   L: N n no not private (0.7) e:hh some beach
2      e:m
3              (1.9)(a)
4   L: are different (0.9)(b) than another
5   T: Uh hh.
6              (*)
7   L: °Than others° .hh and e:m
8              (4.1)(c)
9   L: U:hh .h
10             (2.8)(d)
11  L: Uh
12             (4.2)(e)
13  L: A:nd the beach .h e:hh intensive
14     tourists
15             (1.7)
16  T: °a lot of tourists°=
17  L: =°a lot of tourists° .h[h   e]:hh they
18  T:                          [hm mm]

19  L: (0.6) they can do easily
```

The frequency of hesitation markers in the learner's talk displays uncertainty about the coming talk. There are pauses and a marked withholding of help from the teacher, e.g.pauses (a) to (e) are potential sites where T could have provided affiliative talk or assistance. This lack of talk signals further work by L is required before alignment (Tarplee 1993). Note that in line 5, T does provide a minimal affiliative receipt, "Uh hh", but responsibility for speakership remains with L. (Schegloff

1982). The learner actions a self-repair in line 7. The learner's turn, lines 13-14, includes the repairable 'intensive'. A (1.7) pause follows representing an opportunity point for learner self-repair or repair-initiation. However, there is no display made of awareness of error or any repair attempts from L. The teacher actions a correction. The repairable is picked out and is redone as "a lot of tourists". In this correction, a) there are no explicit repair markers, b) no surrounding syntactic frame, c) no stress pattern to highlight the repair, d) an even intonation, e) it is quieter than the surrounding talk, and f) it is imitated by the learner in receipt, this imitation is pitch-matched. The repair is attended to by teacher and learner in a minimalistic way and does not become the focus of the talk. The learner does an imitation/redoing of the repair in line 17 and makes a claim for continuing speakership, ".hh e:hh they (0.6)". The teacher does a minimal receipt of the learner's redoing in overlap with this claim and also signals the learner's responsibility for continuing the talk, "hm mm" in line 18 (Schegloff 1982) In contrast to extract 1, the 'camouflaged' other-correction in this extract has economically and swiftly dealt with the need for repair and avoided potentially lengthy repair-initiation which could provide further problematic talk. The agenda of this lesson, in contrast to ZLI:SFM:C1, is creating and getting on with conversation and this is indexed in the design of the talk. Exposed and explicit forms of repair would have had a different interactional cost. Consider extract 3 below which demonstrates further camouflaging characteristics.

Extract 3: ZLI:SFM:GB1

```
1   L:  A hat (.) u::h is belong- a hat
2                (1.0)
3   L:  Is belong
4                (4.0)
5   L:  Yes (.) to Gre- Greece.
6                (1.0)
7   T:  So the hat comes from (.) Greece..
8   L:  Yes Greece..
9   T:  °Yes°.
10  L:  Greece and e:hm
```

11 (2.0)
12 L: °Black°
13 (1.2)
14 L: °Clothes°
15 (1.0)
16 °Comes from°
17 (1.0)
18 L: E::i ehh (*) A- Africa.
19 T: °Right°=
20 L: =°Africa°.

The hesitancy, cut-offs in the learner's turns and pauses signal concern with the coming talk. The teacher refrains from assisting in spite of the various pause opportunities. The learner makes another attempt at completing her turn in 3. No assistance is requested from the teacher and none is offered. There is also a lack of affiliative talk from the teacher; no 'yes' or minimal 'hm' receipts. This lack of affiliation signals that further work is required (Tarplee 1993). However, after a 4.0 pause the learner explicitly displays her own assessment of her talk and she then completes her turn. A 1.0 pause follows and the teacher provides an upshot, a clarification request, of the learner's prior talk in line 7. The upshot a) displays, to the learner, the teacher's understanding of her talk, b) summarises the prior talk, c) projects the opportunity for learner alignment, or non-alignment which would project potential further work is necessary before affiliation, and d) is a candidate model. The learner does not action a redoing of the repair, but orients to the request for clarification by providing agreement (in line 8). Notice that it is not the specific repair element in this upshot that is intonationally highlighted in the teacher's talk; "So the hat comes from (.) Greece". The focus on the repair activity is therefore downgraded. Evidence to support that L has treated the teacher's talk as a repair is found later in line 16 where the repair is embedded into the learner's talk. The teacher's model is redone, but it is grammatically incorrect in this context.

In the following extract the learner requests help from the teacher and states the nature of the required assistance.

Extract 4: ZLI:SFM:GB1

1 L: last year u:hh (1.0) pt .hh there was a
2 Turkish (1.0) Turkish woman (.) on the beach
3 (3.0)
4 L: Very old and fat
5 (2.0)
6 L: .h he heh an e::h without ((gestures around
7 chest))
8 T: °A bikini top°
9 L °A bikini top°
10 T: °Hm mm°
11 L: I- I'twas horrible

The repair in this fragment comes after learner request for assistance and thus an explicit display of lack of knowledge is made. In line 6 the learner pinpoints the target item with a gesture. The teacher's following repair is isolated from a surrounding syntactic context and is quieter than the surrounding talk. The repair is redone by the learner, it is also quieter than the surrounding talk and is pitch-matched. The teacher follows this ultimate learner self-repair with a minimal receipt which displays that the repair activity has terminated successfully, that no accounting is required and signals the learner's responsibility for on-going speakership.

Extracts 5 and 6 are also taken from a lesson where conversation is the global agenda, but target language has been specified for use. At the beginning of the lesson T has introduced several target phrases. In the extract below the learner requests assistance and the teacher actions a camouflaged repair. The learner's redoing is in overlap with the teacher's repair turn and further working on talk is necessitated in later turns. Repair is made the explicit focus of the talk.

Extract 5: ZLI:SFM:P1

```
1   L:  =failure is (0.1) u:m (0.4) failure is
2       .hh I: think that is somesing (0.4) mm:
3       u:m somesing like what uh like um::: .huh
4                (5.3)
5   L:  like I want to:
6                (2.2)
7   L:  to win (0.3) uh::
8                (1.0)
9   L:  a business and I I I I- and my- and the
10      conqueries- conquerency?
11  T:  competi-tors
12  L:           =competit- competitance uhh
13      (cough) uh
14               (2.0)
15  L:  could uh maybe (0.1) better than me
16               (1.0)
17  T:  okay .hh so (*) failure is perhaps the
18      opposite of success
19  L:  yes (0.1) yes
20  T:  the opposite =of success
21  L:                =yes
22  L:  yes
23               (0.4)
24  T:  okay yes remember the word competitors
25               (0.2)
26  T:   [competitors
27  L:   [competitors
28  T:  y[es
29  L:   [competitors
```

This extract demonstrates how both teacher and learner may control the extent of focus on target language form and thus cost to the interaction. The learner's turns (lines 1-8 incorporate hesitation and pauses. The teacher withholds from assisting or affiliating talk and so leaves responsibility of speakership with the learner. In line 10 the learner

BEST COPY AVAILABLE

displays awareness of a potential problem with his talk, and also that he is unable to execute a repair by himself. L offers two possibilities, the second of which, (marked by question intonation), is oriented to by the teacher as a request for help and repair. The learner's request for help in line 10 is a minimally designed request from the learner and so in itself preserves the focus on topic rather than projecting a detailed digression towards corrective exchanges and explanation of the form of the language. The teacher's other-correction in line 11 also takes a minimal form as it attends to a recent correctable part of the learner's utterance and does it as a single lexical item. The activity of correction is downgraded by both participants. The teacher's repair has no explicit markers, is not embedded in a surrounding syntactic frame, is not highlighted prosodically and is imitated in receipt by the learner. However, on this occasion the learner does the redoing of the repair in overlap with the teacher's repair. The learner's redoing is incorrect, it is not an imitation of the teacher's model. At this point in the talk the learner is not brought to account by the teacher. The talk continues and the learner completes his specific, local goal at this juncture of the lesson; defining the word 'success'. In lines 17-18 the teacher does an upshot of the prior talk. The upshot, as in extract above a) provides an opportunity for learner alignment, b) displays the state of the teacher's understanding of the talk, c) projects an opportunity for further work to be accomplished if affiliation is not accomplished d) models a candidate target for the learner and so assists in the establishment of mutual comprehension between the participants. The learner provides agreement to the teacher's upshot. The teacher follows this with a redoing of part of her upshotting turn. The learner actions further affiliative talk. After the establishment of understanding, the teacher actions an explicit repair of the repairable "competit competitance" as the previous downgraded repair attempt failed and so correction is made the interactional focus. The teacher models the repair once again and this is imitated by the learner. The learner's redoing this time is acknowledged as being acceptable by the teacher with a 'yes' receipt in line 27.

In extract 6, below, the learner displays his inability to action a self-repair. After the teacher's camouflaged repair the learner pursues the correction activity because the repair is not the category he requires.

Extract 6: ZLI:SFM:P1

1 L: look uh an uh (*) my company hadn't uh
2 hadn't uh:m subside o:r subside I don't
3 know
4 T: subsidised
5 L: subsidised subsidised
6 T: hm mm
7 L: subsidised but uh .h what a subsidise u:h
8 T: subsidy
9 L: a subsidy
10 T: subsidy
11 L: uh: subsidy of (*) EC o:r government

The learner explicitly displays that he is not sure about the word he wants (lines 2-3) and is not able to come to a decision about it himself. The teacher's other-correction takes a minimal form; there are no repair markers, no syntactic frame, and it is not highlighted prosodically and is imitated by the learner in receipt. The repair sequence is closed, as in Example A and extract 2 with a minimal "Hm mm" which signals the end of the repair activity, its successful accomplishment and that the learner has responsibility for continuing speakership. However on this occasion the learner is aware that the teacher's correction is not actually what he was searching for and the focus on the form of the language is maintained by the learner. The learner clearly signals the category of the repair that is being requested (in line 7); a noun is required rather than the verb form that was offered by T. This is evidence of real collaboration in repair between T and L. The teacher provides the required repair that has been explicitly sought for by the learner. The repair takes a minimal form once again. The repair is imitated by the learner and his turn proceeds. The teacher keeps the activity of correction to a minimum, whilst the learner who is in possession of sufficient knowledge is able to collaborate in this repair trajectory and maintain focus on the form of the language until the repair is successfully completed.

Extract 7 below illustrates the potential cost of repair initiation to the interaction, lesson agenda and language learner. For comparison, example E below (Jefferson 1987) shows that between participants who share native-speaker competencies there may be little cost to the ongoing interaction. After a potential site for self-repair, (pause in 4), Louise initiates repair by identifying the trouble-source by repeating the repairable (line 5) with rising ('question') intonation. The beginning of the repairable is emphasised by stress, thus locating and marking the repairable. This initiation leads to a self-repair from Ken without delay. Ken overtly marks out the repair with stress. The extent to which the repair takes over the focus of the interaction is kept to a minimum, but both parties highlight their parts of the repair activity.

(Example E)

1 Ken: Hey (.) the first ti:me they
2 stopped me from selling cigarettes
3 was this morning.
4 (1.0)
5 Louise: From selling cigarettes?
6 Ken: Or buying cigarettes.

Extract 7, taken from a lesson where teacher and learner are holding a discussion about topics such as television, books, actresses etc., illustrates the potential cost of repair to the interaction, lesson agenda and language learner. The language work accomplished in the sequence of talk in the extract above does not remain restricted to the replacement of one specific lexical item but is widened to include the displaying of grammatical and syntactic knowledge (concerning the use of 'since', 'for' and 'ago' when referring to points in the past).Therefore there are a number of potential acceptable repairs.

Extract 7: ZLI:DC:G1

L: I: u:m (0.4) pt read something about her an
interview last time I w-was here (0.2) in
London an:d she got oscars already and
since (0.2) two or three (0.1) years she
is a member of (0.2) parliament (0.2)
T: S[:ince]
L: [she be]
T: Since two or three yea:rs,
L: She: (0.1) since two or three years (0.4)
she has been
(0.3
T: No [stop] that was okay but y- b- sin:ce=
L: [She]
(0.2)
T: Two or three years
(0.2)
L: Since two or three ye:ar (0.4) she: has
been
(1.1)
T: (no re-) remember we wrote it=
L: =Hm: since two or [thr- (*)-
[teacher writes on board-
L: Oh no for two or three years s:- sh: she
has been or is (.) uh?
T: >She has been<
L: Has been .h for two or three years she
has been a member of parliament [h]=
T: [°Righ°]
L: =and she belongs to the labour party
(0.2)
T: Or if you use since you could say (0.1) she
h[as been
L: [Sin:ce
(0.2)
T: Since=

```
36  L:  =Si:nce=
37  T   =Two years
38            (1.1)
39  L:  She has been=
40  T:  =s-heh-ince two y-heh-ears
41            (1.0)
42  L:  °Since° (*) °two° (*) years ago
43  T:  Yeh (0.1) yeah cause then y- [you're
44  L:                               [hm
45  T:  fixing it
46  L:  Hm:[m hm since two years ago she has been
47  T:     [ye
48      a member of parliament
```

The teacher attempts a repair initiation in line 6 which pinpoints the site of the repair "s:ince". The initiation fails to generate a successful repair from the learner who does a redoing of his previous talk. The learner proves unable to locate and action a repair based on T's repair initiation. The teacher withholds actioning other-correction and pursues further repair-initiation. T indicates that the talk redone by the learner is not problematic, hence the repairable is located elsewhere. In line 12 the teacher tries to initiate learner self-repair with a reiteration of the repairable 'since' again. The repairable is highlighted by greater stress on this occasion. The learner fails to action a self-repair. Later the teacher alludes to his assumption and belief that the learner is in possession of the knowledge about the target language under focus in this repair sequence as they have worked on this aspect previously; "remember we wrote it" (line 20). The learner is able to action a self-repair and overtly marks his recognition of the repair and realisation of the repair expectations by emphasising the repair element "for" in line? L continues with the local task of finishing the target sentence completion. However the attempt terminates with a quick request for help "uh?" (in line 24). An other-repair is actioned by T. The repair is isolated, but the speed of delivery is increased. The learner does a redoing of part of the teacher's model and after an in-breath does a redoing of the whole target sentence. The focus of the talk on repair and the form of the target language does not finish at this point. In line 31 the T sets up another sentence completion task for the learner but fails to generate an immediate successful learner repair. The repair is accomplished by the learner 11 lines later after repeated initiation attempts. The learner explicitly acknowledges the repair activity as the repairable is marked by stress ("ago" in line 42). The display of lack of knowledge in the learner's turns and failure to identify the repairable and complete a learner self-repair resulted in elongated initiation from T and several failed repair attempts by L. The pursuit of self-repair and withholding of other-correction in this extract ensured that repair became the local agenda and that the learner was forced to display his level of knowledge about a particular aspect of the target language. What happens in extract 7 clearly contrast with repair trajectories where camouflaged other-correction ensured that the ongoing interaction was minimally interrupted. The fact that the teacher had a basis for assuming the level of learner knowledge was alluded to in the talk and may explain his insistence on repair-initiation. Moreover, the repair required more than the replacement of a single lexical item.

Extract 7: ZLI:DC:G1

1 T: So it's difficult
2 L: It was (*) difficult=yes but I understood
3 it because I saw the musical
4 (*)
5 T: Because you saw the musical (*) or because
6 L: I (*) had seen
7 (*)
8 L: Had seen?
9 T: Yeah
10 L: I had seen the musical=
11 T: =Right if you hadn't seen the musical
12 L: I wouldn't=more difficult to understand
13 (*)
14 T: °Right°

The repairable "saw" occurs in (line 3). The learner makes no display of need for repair etc. After a pause (untimed) the teacher initiates repair.

He repeats part of L's prior talk, as in Example E and extract 7 above. The repair is followed by another pause. No repair is attempted by L. T then indicates the site of the repairable in line 5 with a sentence completion task. The learner actions a self-repair. The learner's talk displays uncertainty, a pause in line 6 mid-repair. The lack of affiliative talk from the teacher is oriented to by the learner as a display of a need for further work (Tarplee 1993). The learner does a redoing of the repair with question intonation displaying his uncertainty, but offers no other alternative repairs. The teacher provides affiliative talk in next-turn and maintains the focus on the form of the talk by constructing a sentence completion task which is successfully actioned by L.

Extracts 9 and 10 are from a lesson where correction is the concern of the talk. The teacher and learner are going through sentences written as a homework task. Focus on the form of the target language is an explicit pedagogical agenda in the lesson.

Extract 9: ZLI:A:L1

```
1   L: Yesterday I kept witing do:wn my notes on
2      my carnet °un carnet u:h [I -don't know°]=
3   T:                          [no      n:    ]
4   T: =Note?
5            (0.7)
6   T: Notebook
7            (0.4)
8   L: Notebook
9   T: =Notebook
10           (6.0)
11  T: Right?
```

The lesson activity concerns going through and correcting the learner's homework. The learner's task was to write sentences using specified new language that he has learned on the course. The learner reads out one of his answers (lines 1-2) and explicitly displays that he does not know the word in English that he needs to complete his sentence. The teacher makes repair attempts, which end in cut-offs, in overlap with L's turn. In line 4 the teacher constructs a repair-initiation as a word completion task which fails to engender a learner self-repair. The completion task in itself promotes the activity as a collaborative enterprise. A 0.7 pause follows this initiation attempt and the teacher actions the projected repair; the learner's absence of talk signalling his inability to perform a repair. The teacher's repair is isolated, i.e. without any surrounding syntactic context, as were repairs dealing with the replacement of specific and single lexical items in the learner's talk as in extracts 2, 4, 5 and 6. The repair in extract 9 also generates an imitation by the learner. A difference is that the teacher's repair is highlighted intonationally. Focusing on the form of the language and correction comprise the activity of the talk displayed in extract 9.

In the last extract 10 below, there is more than one source of trouble in the learner's talk. This example is again taken from lesson ZLI:A:L1, where the activity of the talk concerns displaying competency and linguistic knowledge. Lengthened repair initiation , explicit focus on language form and the use of metalanguage characterise the talk as correction is an explicit agenda.

Extract 10: ZLI:A:L1

1 L: Are you sure we go to the wright die- di-
2 uh direction
3 (.) (a)
4 T: °Okay° .hh not we go: (.) (b) h imagine you're in
5 the situation
6 (0.7)
7 L: Uh we ri(de) -°no°
8 T: -Yeh bu- imagine=it's the tense:e
9 (0.4)
10 T: °Lori° =imagine it's now
11 L: Okay
12 (0.7)
13 T: Whi[ch tense would you] use=
14 L: [Are you sure]
15 L: =We are going
16 T: Aright .hh okay an we are going=not to
17 (1.0)

18 T: Not the preposition is no[t to
19 L: [i:n the
20 T: Yes so say it again
21 L: Okay
22 (0.9)
23 T: Say the sentence again
24 L: Alors are you sure we are going in the right
25 de- direction
26 T: Yeh .hh i- uh Lori just say this .h are you
27 sure?
28 (0.8)
29 L: Yes
30 T: Stress the word sure
31 (0.5)
32 L: Are you sure?
33 T: Are you sure (*) we're going
34 (0.4)
35 L: In the wright direction
36 T: In the right direction

The learner reads out his sentence attempt containing the repairables, "go" and "to" in lines 1-2. After a micro-pause, at (a), signalling a coming dispreferred activity, the teacher receipts the turn and then actions a repair-initiation. The initiation identifies one of the trouble-sources. A micro-pause follows at (b) and the teacher provides further initiation, a "cluing" (McHoul 1990). After a 0.7 pause the learner attempts a repair but rejects his repair himself. The teacher withholds from other-correction and pursues further initiation. T explicitly states that the learner has used the wrong tense. The teacher provides two further initiations in lines 10 and 13 before the learner actions a self-repair. T receipts the learner repair in line 16. The teacher then directly proceeds to attend to a second repairable. The teacher's first initiation is minimally packaged and identifies the site of trouble, "not to". There is a one second interval and T continues with further initiation, avoiding other-correction. T highlights the repairable again. The learner actions a self-repair (line 19) and is requested to do a redoing of the repaired stretch of talk (line 20). The activity of the talk now turns to pronunciation business with a sequence in which the talk focuses on intonation and stress.

The nature of the activity of the talk in this extract concerned overt focus on language form and correctness. The lengthened repair initiation sequence ensured that correction remained the explicit business.

6. Concluding remarks

The CA analysis of repair in EFL classroom talk reported in this paper gives testament to the nature of the joint management of issues related to second language development; issues connected with intelligibility, repairing troubles and establishing mutual comprehensibility and intersubjectivity. The description of one of the chief enterprises in EFL classroom talk generated by this CA analysis, is vastly different from the view of reactionary correction and appraisal, typified by 'initiation-response-feedback' routines, deemed to be paradigmatic of classroom talk (Sinclair and Coulthard 1975). Rather than segmenting EFL conversation into such uni-directional categories as initiation, response, teacher negative feedback, etc, correction, as part of the broader phenomenon of repair, has been revealed as an activity which is negotiated by EFL participants on a turn-by-turn basis as they collaboratively work on the re-construction of their talk.

Repair strategies have been shown to impose different costs on the lesson agenda and the learners. Teachers have also been seen to orient to the status of other-correction as a dispreferred activity, by a), restraining from other-correction, b), pursuing repair initiation to increase opportunities for self-repair, and c), packaging other-correction when actioned in an accommodating, 'camouflaged', (e.g. isolation of the repair, delivered at a volume which is quieter than the surrounding talk, and lack of intonational marking), environment which serves to tone down unmodulated other-correction and take the focus off the activity of repair. The 'camouflaged' corrections empowered the EFL teacher to attend to the repair of trouble-sources, but did not oblige a lengthened, explicit or consciously motivated focus on language form. As an example, extract 6, demonstrated that where the L2 learner is in possession of the necessary knowledge he/she may accept the correction in an exposed receipt and even make the correction the focus of the talk him/herself. Repair and control of preference organisation is potentially actionable by both teacher and learner and is negotiated on a 'here and now' basis as their talk unfolds. For example, where the learner displays no awareness of error or inability to action self-repair in their turns-at-talk the EFL teacher may action other-correction in either an exposed or embedded form. What is projected as a relevant next is therefore controlled, to some extent or other, by the teacher and (subject to his/her level of competence) the learner.

Forms of correction were shown to orient to the pedagogic goal of the type of EFL lesson or activity in an EFL class which entails the conscious analysis of aspects of the target language, e.g. a grammar lesson, as in extract 1, 'correcting homework', as in extracts 9 and 10. These types of teaching agendas contrast with lessons or activities in which conversational practice is the global pedagogic goal, as in the discussions of extracts 2, 3, and 4. Explicit forms of correction and their accompanying accountings would require an investment in the talk and make demands on the learner which could prove to be beyond their level of competence. The extended repair activities of extracts 5 and 7 are examples where local agendas become relevant as the talk proceeds and so correction becomes the overt activity of the talk. In extract 5 the teacher actions explicit repair after a 'camouflaged' attempt failed. In extract 7 the teacher displays that he has good reason to anticipate the learner's capacity for self-repair.

This paper has examined the organisational devices which provide for flexibility, local-management and negotiation in the accomplishment of immediate and global interactional agendas in EFL classroom talk.

REFERENCES

Allwright, R.L. and K. M. Bailey. (1991). *Focus on the Language Classroom*. Cambridge: Cambridge University Press.

Iles, Z. L. (1995). 'Learner control in repair in the EFL classroom'. Paper to be presented at BAAL Annual Meeting, September 1995.

Jefferson, G. (1987). 'On exposed and embedded correction in conversation'. In Button, G., and J. R. Lee (eds), *Talk and Social Organisation*. Clevedon: Multilingual Matters. 86-100.

McHoul, A. (1990).'The organisation of repair in classroom talk'. *Language in Society*. 19. 349-377.

Lennon, P. (1991). Error and the very advanced learner, *IRAL*. Vol XXIX/1. 30-44.

Pica, T. (1994). Questions from the language classroom: Research Perspectives. *TESOL Quarterly*. Vol. 28/1. 49-79

Schegloff, E.A. (1982). Discourse as an interactional achievement: Some issues of 'uh huh' and other things that come between sentences. In D. Tannen (ed.) *Analyzing Discourse: Text and Talk*. Washington D.C: Georgetown U.P. 71-93.

Schegloff, E.A. (1992). In another context. In A. Duranti and C. Goodwin (eds.) *Rethinking context: language as an interactive phenomenon*. Cambridge: Cambridge University Press. 191-227.

Schegloff, E.A., Jefferson, G. and H. Sacks. (1977). The preference for self-correction in the organization of repair in conversation. *Language*. 53. 361-382.

Sinclair, J. and M. Coulthard. (1975). *Towards an analysis of discourse*. Oxford: Oxford University Press.

Tarplee, C. (1993). 'Working on Talk: the collaborative shaping of linguistic skills within child-adult interaction'. *Unpublished DPhil thesis*. University of York.

A TIMING MODEL FOR FAST FRENCH*

Eric Keller and Brigitte Zellner
University of Lausanne

1. Introduction

Previous research on the prediction of speech timing has documented influences at three major levels: the phoneme or segmental, the syllabic and the phrase level. In this paper we describe a three-tiered statistical model which has been created for predicting the temporal structure of French, as produced by a single, highly fluent speaker at a fast speech rate. The first tier models segmental influences due to phoneme type and contextual interactions between phoneme types. The second tier models syllable-level influences of lexical *vs.* grammatical status of the containing word, presence of schwa and the position within the word. The third tier models utterance-final lengthening. The output of the complete model correlates with the original corpus of 1204 syllables at an overall $r = 0.846$. However, an examination of subsets of the complete data set revealed considerable variation in the closeness of fit of the model. Residuals have a normal distribution.

1.1. Models Based on the Prediction of Segmental Durations

The most influential statistical model for spoken French text has probably been the model proposed by O'Shaughnessy (1981, 1984). On the basis of numerous readings of a short text containing all phonemes of French, a model of durations of acoustic segments suitable for synthesis by rule was proposed. In this model, 33 rules for the modification of segment duration according to segment type, segment

---

* **Authors' address for correspondence:** Laboratoire d'analyse informatique de la parole (LAIP). Informatique — Lettres, Université de Lausanne, CH-1015 LAUSANNE, Switzerland.

York Papers in Linguistics 17 (1996) 53-75
© Eric Keller and Brigitte Zellner

position and phoneme context served to specify basic phoneme durations.

For sound classes that did not involve prepausal lengthening, the model was able to predict the durations for 281 segments of a text with a standard deviation of 9 ms. But it was less accurate for the prediction of prepausal vowel durations, because of the greater variability of segments in such positions. Moreover, this model was not able to predict silent inter-lexical pauses.

O'Shaughnessy's statistical model is constructed around the hypothesis that speech timing phenomena can be captured by the segment, as if this unit "possesses an inherent target value in terms of articulation or acoustic manifestation" (Fujimura 1981). However, recent measures have indicated that syllable-sized durations are generally less variable than subsyllabic durations, and thus may represent more reliable anchor points for the calculation of a general timing structure than segmental durations (Barbosa and Bailly 1993; Keller 1993; Zellner 1994). The taking into account of explicit syllable-level information is further supported by the observation that stress variations and variations of speech rate tend to modify at least syllable-sized units.

Bartkova's model (1985, 1991) attempts to solve these deficiencies by adding calculated coefficients to the formula for predicting segment durations:

$$Dur\ Seg = DurI + k_{Syll} + k_{Ac}$$

where $DurI$ is the intrinsic duration of the segment, k_{Syll} is a syllabic coefficient, and k_{Ac} an accentuation coefficient. The exact manner in which these coefficients are obtained is not described; it is only noticed that they can vary from a minimum to a maximum interval, according to the position of the segment in the speech chain, and according to the acoustic properties of the speech sound.

The syllabic coefficient depends on the nature of the word (lexical/grammatical), and on the position in the word (initial, medial, final syllable). The coefficient of accentuation depends on the next consonant, on the presence/absence of a syntactic boundary in the case of a final vowel, or on the presence/absence of clusters in the case of a final consonant, as well as on the syllabic structure near a pause.

According to Bartkova, a comparison of predicted and measured durations in 10 sentences gives rather good predictions, since the mean difference on segmental duration is about ±15 ms.

However, it would seem that beyond the opacity of the coefficients, a divergence between predicted and measured durations of the order of 15 to 30 ms can be a major handicap for short segments. In our corpus, for example, the mean duration for /d/ was 50 ms. In the case of such a short phoneme, a 15-30 ms divergence would correspond to an error of 30-60% with respect to its measured duration.

1.2. Required Macro-timing Information

Since the segmental unit cannot capture the overall temporal structure of speech, the next level which can be expected to encapsulate temporal phenomena is the syllable. This appears to be a good candidate. According to some psycholinguists, it is considered to be the minimal perception unit, and according to a number of phoneticians and phonologists, it is the minimal unit of rhythm (see Delais 1994).

It has been shown that quite a number of parameters are involved in variations of syllabic duration. The most important are: the position in the prosodic group, the position in the word, degree of stress, the length of the prosodic group, the position according to the stressed syllable, the position according to the local speech rate (as measured by cycles of speeding up and slowing down), semantic focus, proximity of syntactic boundaries, the status of the word (lexical or grammatical), and emotional factors (Bartkova 1985, 1992; Campbell 1992; Delais 1994; Duez, 1985, 1987; Fant and al. 1991; Fònagy 1992; Grégoire 1899; Grosjean et al. 1975, 1983; Guaïtella 1992; Konopczynski 1986; Martin 1987; Mertens 1987; Monnin et al. 1993; Pasdeloup 1988, 1990, 1992; Wenk et al. 1982; Wunderli. 1987). Some of these factors may be redundant; for instance, in many cases of read text, lexeme-final position may be redundant with phrase-final position.

In view of existing information, it thus seems best to begin with segmental predictions, and to consider syllabic information as additional information which is not captured at the segmental level. One of the important points to consider in the present study will be the selection of non-redundant and relevant information.

Beyond the syllabic level, it is likely that a good predictive model will eventually need to incorporate further information at the word or the phrase level. For example, the prediction of pauses for slow speech requires phrasal knowledge, which is not captured at the segmental or at the syllabic level. In the area of word group boundaries in French speech, a great deal of work has been accomplished to determine the nature of these groups — syntactic groups, prosodic groups, rythmic groups, intonational groups, the congruence between these labels — and to calculate the automatic generation of such groups and potential inter-group pauses (Delais. 1994; Grosjean et al. 1975; Keller et al. 1993; Martin 1987; Monnin et al. 1993; Pasdeloup 1988; Saint-Bonnet et al, 1977). These effects will have to be integrated into a general timing model for a given language, but were not taken into account in the present study.

In the current study, the objective was to account for a single speaker's syllable durations with the smallest number of segmental and syllabic factors. At each succeeding level, relevant parameters were chosen so as to explain the greatest proportion of the variance in the residue of the previous analysis. In this manner, a three-tier model, based successively on segmental, syllabic and phrasal information, was constructed.

2. Method

2.1. The corpus

A highly fluent speaker of French (a professor of French literature) was recorded with 277 sentences, the first 100 of which were analysed for the present study. The speaker was instructed to speak quite rapidly, with a normal, unexaggerated intonation. The resulting readings have generally been judged by listeners as highly intelligible and well-pronounced. No dialectal particularities were noted.

Recording occurred in studio conditions on DAT-tape. The digitized data was transferred to Macintosh computer and was downsampled to 16 kHz.

2.2. Time labelling

The time occupied by each phoneme was labelled with the Signalyze™ program according to detailed instructions on how to handle phoneme-to-phoneme transitions (Thévoz and Enkerli 1994). Specifically, transitions in the acoustic corpus was analyzed according to three articulatory levels: labial, lingual and laryngeal. For example, the coarticulatory overlap at the /e/-/s/ transition was marked by symbols representing the following events: "onset of friction, associated with the lingual level", followed at a given time interval by an "offset of fundamental frequency, associated with a cessation of vocal cord activity". The following possible states were distinguished:

Labial system: aperture, occlusion, friction, burst, error

Lingual system: aperture, occlusion, friction, burst, palatal, transient movement, error

Laryngeal system: aperture, occlusion, transient movement, diminution, error

"Error" refers to any state that occurs inadvertently, such as during a speech error.

To examine the reliability of transcriptions, two judges compared judgements concerning how and where points of transition between inferred articulatory states were to be marked. Two measures of interjudgemental agreement were used:

Robustness (agreement in the application of criteria to state transition), scored 1 = low agreement, 2 = agreement in general, but some further discussion required, and 3 = excellent agreement.

Precision, scored 1 = more than two Fo periods difference, 2 = 1-2 Fo periods difference and 3 = less than 1 Fo period difference in measurement.

Both measures showed good to excellent interjudgemental agreement. Over the 50 types of state transitions examined, there were no cases of low robustness or low precision. The average robustness was 2.53 and the average precision was 2.68.

A total of 4544 phonemes and 1203 syllables were analyzed in this manner.

3. Analysis and Results

A modified step-wise statistical regression technique was used to develop a well-fitting model of this speaker's timing behaviour. In accordance with previous observations on factors that influence speech timing, it was decided to model three major levels: the segmental, the syllabic and the phrase level. In step-wise fashion, each succeeding level was made to model the residue left by the previous level. Three different models were thus established, the Segmental, the Syllabic and the Phrase Model (Figure 1).

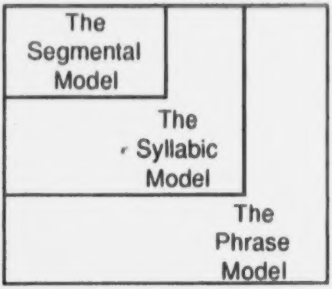


Figure 1. The Segmental, Syllabic and Phrase Models. Each subsequent model incorporates the modelling effects of the previous level.

3.1. Model 1: The Segmental Model

Segmental Durations and Overlap Zones. An initial issue concerned the calculation of segmental duration in a corpus where coarticulatory transition zones are marked explicitly. Does phoneme duration correspond to the zone of the signal which is unambiguously marked for a given phoneme (zone B in figure 2), or does it include one or both zones of coarticulatory overlap with adjoining phonemes (zones A and C in figure 2)?

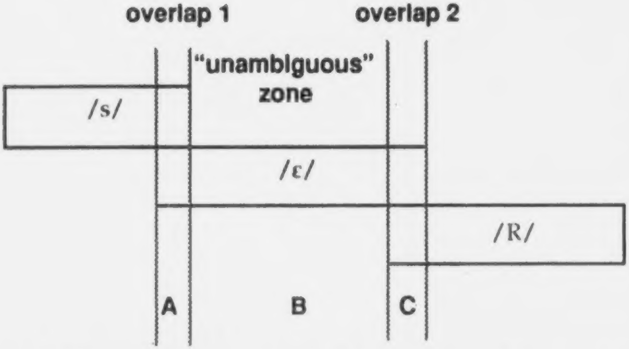


Figure 2. What constitutes a phoneme? B is a portion of the signal that is unambiguously marked for the phoneme /ɛ/, while A and C are transitory zones with adjoining phonemes.

The issue was resolved with reference to durational variation. The combination of zones A, B and C (with an average coefficient of variation of 0.375) turned out to be systematically less variable than the unambiguous zone B (with an average coefficient of variation of 0.412) (see Table 1).

	A	B	C
Average coefficient of variation (s.d./ mean) for 34 phonemes	1.6379	0.4123	1.7472
	A + B	**B + C**	**A + B + C**
Average coefficient of variation for 34 phonemes	0.3916	0.3933	0.3751

Table 1. Coefficients of variation for zones A, B and C as well as various combinations of these zones

Also, combinations of zones A and B, or of B and C, were less variable than zone B alone. The transition zones can thus be considered to be "buffer zones" whose function, in part, may well be to "regularise" phoneme duration. For the purpose of the present research it was thus decided to consider the combined duration of A, B and C as "phoneme duration". Syllable durations were constructed from phoneme durations by taking into account transitional overlaps. As a net effect, the segmental duration entering the statistical modelling procedure is slightly more regular than more commonly measured phoneme durations. Nevertheless, it is not believed that the modelling results of the present study seriously depend on this manner of proceeding; the size and resilience of the measured effects suggest that as long as transitions are handled in systematic fashion, the predictive pattern should remain largely identical.

3.2 Segmental transformation and grouping.

Raw segment durations were non-normal in their distribution. Among the common transformations, the log10 transformation produced the closest approximation to a normal distribution (Figure 3a, b). All calculations of the segmental portion of the model were thus performed on log10-transformed durations.

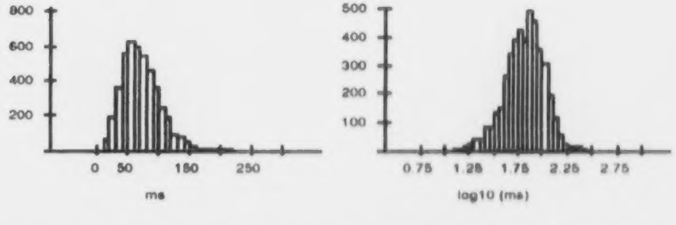


Figure 3a. The distribution of segment durations before and after the log 10 transformation: histograms.

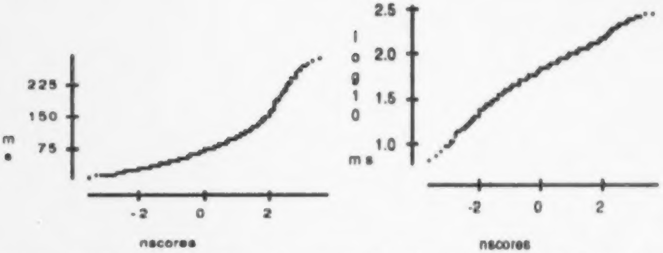


Figure 3b. The distribution of segment durations before and after the log 10 transformation: normal probability plots.

Subsequent to transformation, phonemes were grouped according to their mean durations and their articulatory definitions. Eight classes could be identified (Table 2). Groups showed roughly comparable coefficients of variation, and an inspection of histograms and normal probability plots showed roughly normal distributions for all classes whose *N* was greater than 100.

Phoneme type	Name	Mean duration (ms)
œ, ø	AntRound	109.45
ʃsf	Fric	105.17
œ̃, ɛ̃, ɑ̃, ɔ̃	Nas	97.78
o	PostMidRnd	94.92
p, t, k	UnvPlos	92.94
a, e, ɛ, ɔ, u, i, y	OthVow	69.62
b, z, m, ɲ, g, v, ʒ, n, d, ʔ	VcdCons	61.72
R, j, w, l, ɥ	SemiVLiquids	43.63
Mean		90.23

Table 2. Mean durations for phoneme classes (N = 4544)

Phoneme type	Coefficient of variation (s.d./mean)	Frequency (*N*)
œ, ø	0.4881	71
ʃsf	0.2708	357
œ̃, ɛ̃, ã, õ	0.3585	334
o	0.3130	60
p, t, k	0.3475	504
a, e, ɛ, ɔ, u, i, y	0.4089	1557
b, z, m, ŋ, g, v, ʒ, n, d, ʔ	0.3669	892
R, j, w, l, ɥ	0.4908	769
Mean	0.3648	539

Table 2.(continued) Mean durations for phoneme classes (N = 4544)

To test Model 1 in the syllabic context, square root-transformed syllable durations were calculated on the basis of coefficients produced by the linear model for segmental durations, and by taking into account mean durations of phoneme-to-phoneme transitions. These calculated syllable durations were compared to the square root-transformed measured syllable durations. The correlation coefficient was $r = .647$ ($N = 1203$, $p<.0001$) (Figure 5).

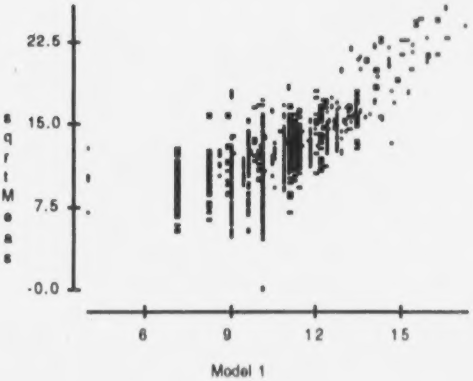


Figure 5. Prediction of the Segmental Model (Model 1): Syllable durations predicted exclusively on the basis of segmental durations (r = .647). Values are in sqrt(ms).

The residue from the model (= observed - predicted) was termed "Delta 1" and served as the basis for further factorial modelling at the syllabic level.

3.3 A Linear Model for Segmental Durations.

Using the Data Desk® statistical package on the Macintosh, a general linear model for discontinuous data (based on an ANOVA) was calculated with partial (non-sequential, Type 3) sums of squares. The following main and interaction factors (up to two-way[1]) were postulated:

duration (log10(ms)) = constant + previous type + current type + next type + previous type * current type + current type * next type + previous type * next type

---

[1] For reasons of insufficiency in per-cell observations, calculation complexity and theoretical difficulty of interpretation, three-way interactions were not calculated.

Table 3. The Segmental Model: Analysis of Variance for Segmental Data (N = 4544) Using Partial Sums of Squares

Source	df	Sums of Squares	Mean Square
Const	1	14903.8	14903.8
previous	8	0.123239	0.015405
current	7	3.13402	0.447717
next	8	0.267002	0.033375
previous * current	50	3.24144	0.064829
current * next	50	5.04499	0.100900
previous * next	60	1.79531	0.029922
Error	4360	101.137	0.023197
Total	4543	196.070	

Source	df	F-ratio	Prob
Const	1	642500	≤ 0.0001
previous	8	0.66410	0.7236
current	7	19.301	≤ 0.0001
next	8	1.4388	0.1748
previous * current	50	2.7948	≤ 0.0001
current * next	50	4.3498	≤ 0.0001
previous * next	60	1.2899	0.0665
Error	4360		
Total	4543		

In the partial sums of squares solution, all factors were significant at $p<.05$, with the exception of "previous type" and "next type", taken alone, and the interaction term "previous type * next type" (Table 3). The residual error was 101.137/196.070 = 0.516, that is, the model explained about 48.4% of the variance. Expressed in terms of a Pearson product-moment correlation, the model's predicted segmental durations correlated with empirical phoneme durations at $r = 0.696$.

3.4 Syllable Durations and Delta 1.

Another means of testing the model is a comparison with measured syllable durations. In contrast to phoneme durations, where a log transformation served to provide roughly normal distributions, square roots had to be applied to measured syllable durations in order to approximate normal distributions (Figure 4).

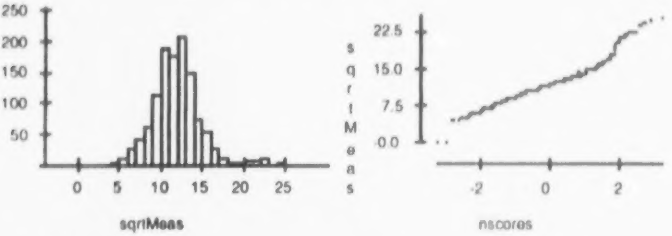


Figure 4. Syllable durations in ms were square-root transformed in order to approximate a normal distribution.

3.4.1. Model 2: The Syllabic Model

Syllabic Factors Predicting Delta 1. After considerable experimentation with a variety of factors described in the literature, a three-factor model, including two-way interactions, was retained for analysis:

delta 1 = constant + function + position + schwa + function * position
+ function * schwa + position * schwa,

where *"function"* distinguishes whether the syllable is found in a lexical or a function word, *"position"* identifies three types of position in the word which are (1) "monosyllabic and polysyllabic-initial", (2) "polysyllabic pre-schwa" and (3) "other", and "schwa" indicates whether or not a schwa is present in the syllable. Again, a general linear model for discontinuous data was calculated with partial (Type 3) sums of squares. The results of the ANOVA showed that all main and interaction factors were significant at $p<.05$ (Table 4). The residual error of 3277.29/5432.93 = .6 indicated that the model explained 40% of the variance in Delta 1.

Table 4. Analysis of Variance for Delta 1 (N = 1203) Using Partial Sums of Squares

Source	df	Sums of Squares	Mean Square
Const	1	2663.53	2663.53
function	1	176.508	176.508
position	2	98.5753	49.2877
schwa	1	149.296	149.296
function * position	2	97.3872	48.6936
function * schwa	1	27.5860	27.5860
position * schwa	2	63.0467	31.5234
Error	1193	3277.29	2.74710
Total	1202	5432.93	

Source	df	F-ratio	Prob
Const	1	969.58	≤ 0.0001
function	1	64.252	≤ 0.0001
position	2	17.942	≤ 0.0001
schwa	1	54.347	≤ 0.0001
function * position	2	17.725	≤ 0.0001
function * schwa	1	10.042	0.0016
position * schwa	2	11.475	≤ 0.0001
Error	1193		
Total	1202		

Model 2 and Delta 2. Syllable durations obtained from the segmental model were combined with those from the present linear model for Delta 1 to produce the Syllabic Model (Model 2). The predictions correlated with observed square root-transformed syllable durations at $r = .723$ (N=1203) (Figure 6). The residual data was termed Delta 2.

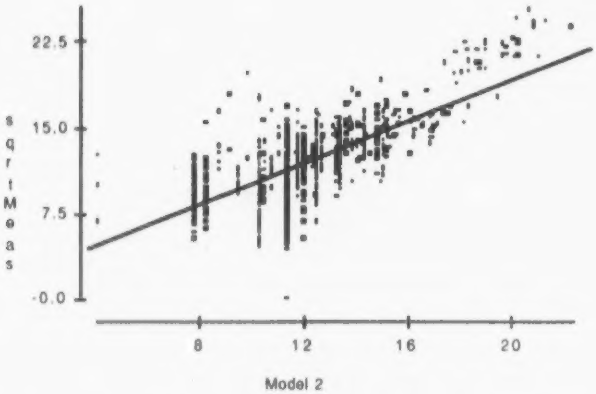


Figure 6. Prediction of the Syllabic Model (Model 2): Syllable durations predicted on the basis of segmental durations and syllable-level factors (r = .723). Values are in sqrt(ms).

3.5. Model 3: The Phrase Model

Inspection of the predictions of Models 1 and 2 (Figures 5 and 6) showed a noticeable deviation from the regression line in the higher values. Specifically, these models underestimated most syllable durations in the > 280 ms range. Furthermore, an examination of Delta 2 revealed that the residual error was most pronounced for utterance-final syllables ending in a consonant. Consequently, a correction term was calculated, which was applied to such syllables in Model 3.

The predictions of Model 3, which incorporates segmental and syllabic modelling as well as the phrase-final correction term, correlated with the observed square root-transformed syllable durations at r = .846 (Figure 7). The residual values from Model 3 vary quasi-randomly around 0. At the present time, it appears that only more sophisticated rules for the generation of the schwa vowel may still be able to improve this model's predictive capacity to some degree.

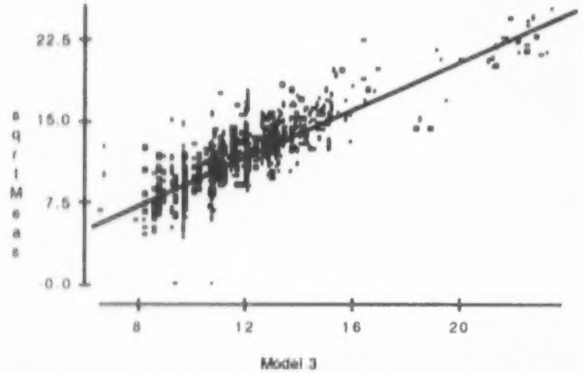


Figure 7. Prediction of the Phrase Model (Model 3): Syllable durations predicted on the basis of segmental durations, syllable-level factors and phrase-final lengthening (r = .846). Values are in sqrt(ms).

3.5.1. Stability

The Phrase Model was examined for its predictive stability by performing Pearson product-moment correlations between various subsamples of the data and the model's prediction. The resulting data is presented in Table 5.

Table 5. Pearson Product-Moment Correlations between Various Subsets of the Dataset and the Phrase Model' s Prediction

	slices of 50 syllables	slices of 100 syllables
1st slice	0.9	0.884
2nd slice	0.87	0.872
3rd slice	0.853	0.852
4th slice	0.89	0.726
5th slice	0.866	0.823
6th slice	0.852	0.868

	slices of 200 syllables	slices of 300 syllables
1st slice	0.878	0.869
2nd slice	0.789	0.805
3rd slice	0.838	0.874
4th slice	0.885	0.838
5th slice	0.841	
6th slice	0.838	

Table 5. (Continued) Pearson Product-Moment Correlations between Various Subsets of the Dataset and the Phrase Model's Prediction

It can be seen that the model's predictive capacity varies considerably from one subset to the next. For example, the correlation was only .726 for the fourth slice of 100 syllables in the set, while it had been .884 for the first slice. Even when slices of 300 syllables are compared, considerable variability prevails. The reasons for these instabilities are presently being investigated.

4. Discussion

By a modified step-wise procedure, a general model for the prediction of the fast-speech performance of a highly fluent speaker of French was constructed. The initial model incorporates segmental information concerning type of phoneme and proximal phonemic context. The subsequent model adds information about whether the syllable occurs in a function or a lexical word, on whether the syllable contains a schwa and on where in the word the syllable is located. The final model adds information on phrase-final lengthening. The effects of these three levels are demonstrated on a single sentence in Figure 8. In view of current discussions surrounding segmental and syllabic contributions to timing models, it is interesting to note that segmental information accounts for a major portion of the variance explained by the model. As Figure 8 shows, segmental information alone successfully predicts several cases of major syllable lengthening.

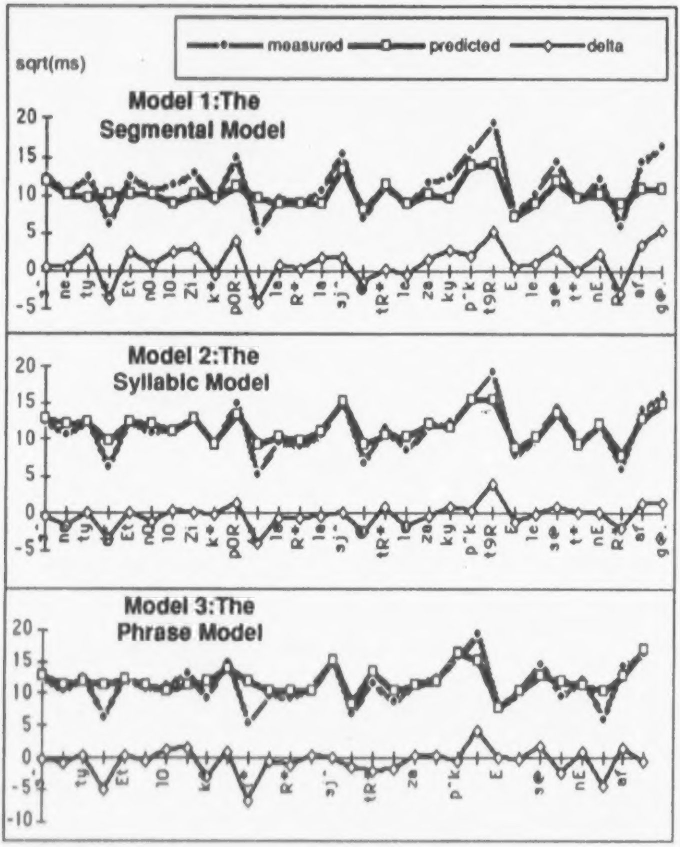


Figure 8. A comparison of predictions of the three models and measured syllable durations for the sentence "Son étude ethnologique porte sur la relation entre les acupuncteurs et les centenaires afghans".

The overall correlation of 0.846 between predictions of Model 3 and the data set from which the model is derived is encouraging. This correlation level corresponds roughly to the average inter-speaker correlation of $r = 0.833$ for phrase-final syllable durations, as measured between the readings of a short text by 12 speakers in the Caelen-Haumont corpus (Caelen-Haumont 1991; see Keller 1994). This means that the model behaves as differently from its target data as one natural speaker would behave with respect to another speaker. Although this may be an acceptable initial predictive level for synthesis purposes, further improvements in the modelling would be welcome. Preliminary indications suggest that such improvements may come about through predictions of the presence vs. the absence of schwa, through explicit predictions of the effects of speech rate manipulation, and in longer texts, through a better modelling of pauses. Further information on possible improvements may also be gained through an examination of cases of high delta 3 values in subsets of the present data set. These effects are currently being studied.

It is worth noting that in the present fast-speech corpus, no phrase-level effects were identified, other than phrase-final lengthening. This is in contrast to our findings on the production of French at a normal speech rate, where a fairly systematic increase of lexeme-final syllable durations was observed over the extent of the prosodic phrase (Keller *et al.*. 1993). It seems likely that in conditions of considerably accelerated speech rate, our speaker sacrificed some of the "niceties" of phrase-internal timing modulation, and limited himself to a single, phrase-final durational marker.

Considerably more work also needs to be done before the generalisability of the present model can be tested. The examination of the model's stability has shown that predictions begin to show comparable strength at about 300 syllables or more. Consequently, systematic testing of these predictions for another speaker would involve a completely new research study. Nevertheless, a few quick examinations of predictions for another speaker's sentences suggest that the model may indeed be generalisable to more than one speaker of French (Figure 9)[2].

---

[2] The authors are grateful to the following members of the LAIP team for their invaluable assistance in scoring and creating the present corpus: Nicolas Thévoz, Alexandre Enkerli, Hervé Mesot, Cédric Bourquart, Nicole Blanchoud, and Thomas Styger. Particular thanks go to Prof. J. Local (York

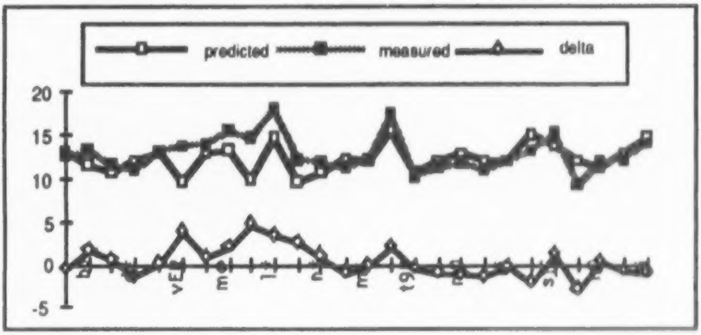


Figure 9. A comparison of predictions of Model 3 and the measured syllable durations of another speaker of French for the fast reading of the sentence "Beaucoup de gouvernements voient le CERN comme un moteur de modernisation technologique".

REFERENCES

Barbosa, P. and Bailly, G. (1993). Generation and evaluation of rhythmic patterns for text-to-speech synthesis. *Proceedings of ESCA Workshop on Prosody*. Lund: Sweden. 66-69.

Bartkova, K. (1985). Nouvelle approche dans le modèle de prédiction de la durée segmentale. *14ème JEP*. Paris. 188-191.

University, UK) for his many ideas and his encouragement. Prof. A. Wyss of the University of Lausanne is cordially thanked for his participation as a subject for this study. This research is supported by the Fonds National de Recherches Suisses (Projet Prioritaire en informatique and ESPRIT Speech Maps) and by the Office Fédéral pour l'Education et la Science (COST-233).

Bartkova, K. (1991). Speaking rate in French application to speech synthesis. *XIIème Congrès International des Sciences Phonétiques, Aix en Provence. Actes.* 482-485.

Caelen-Haumont, G. (1991). *Stratégies des locuteurs et consignes de lecture d'un texte: Analyse des interactions entre modèles syntaxiques, sémantiques, pragmatique et paramètres prosodiques,* Thèse d'Etat, Aix-en-Provence.

Campbell, W.N. (1992). Syllable-based segmental duration. *Talking Machines. Theories, Models, and Designs.* Amsterdam: Elsevier Science Publishers. 211-224.

Delais, E. (1994). Prédiction de la variabilité dans la distribution des accents et les découpages prosodiques en français. *XXèmes Journées d'Etude sur la Parole* . Trégastel. 379-384.

Delais, E. (1994a). Rythme et structure prosodique en Français. In C. Lyche (ed). *French generatve phonology: retrospective and perspectives.* Eurpean Studies Research Institue, Salford. 131-150.

Duez, D. and Nishinuma, Y. (1987). Vitesse d'élocution et durée des syllabes et de leurs constituants en français parlé. *Travaux de l'Institut de Phonétique d'Aix.* 11. 157-180.

Duez, D., Nishinuma, Y. (1985). Le rythme en français. *Travaux de l'Institut de Phonétique d'Aix.* 10. 151-169

Fant, G., Kruckenberg, A. and Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics.* 19. 351-365.

Fònagy, I. (1992). Fonctions de la durée vocalique. In P. Martin (Ed.), *Mélanges Léon.* Editions Mélodie-Toronto. 141-164.

Fujimura, O. (1981). Temporal organisation of articulatory movements as a multidimensional phrasal structure. *Phonetica.* 38. 66-83.

Grégoire, A. (1899). Variation de la durée de la syllabe en français. *La Parole.* 1. 161-176.

Grosjean, F. (1983). How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics.* 21. 501-529.

Grosjean, F., and Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français. *Phonetica.* 31. 144-184.

Keller, E. (1993). *Prosodic Processing for TTS Systems: Durational Prediction in English Suprasegmentals.* Final Report, Fellowship, British Telecom.

BEST COPY AVAILABLE

Keller, E., Zellner, B., Werner, S., and Blanchoud, N. (1993). The Prediction of Prosodic Timing: Rules for Final Syllable Lengthening in French. *Proceedings,of ESCA Workshop on Prosody*. Lund, Sweden. 212-215.

Keller, E. (1994). Fundamentals of phonetic science. In E. Keller (ed.), *Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art and Future Challenges*. Chichester: John Wiley. 5-21.

Konopczynski, G. (1986). Vers un modèle développemental du rythme français: Problèmes d'isochronie reconsidérés à la lumière des données de l'acquisition du langage. *Bulletin de l'Institut de Phonétique de Grenoble*. 15. 157-190.

Martin, P. (1987). Structure rythmique de la phrase française. Statut théorique et données expérimentales. *Proceedings des 16e JEP. Hammamet*. 255-257.

Mertens, P. (1987). L'intonation du français. De la description linguistique à la reconnaissance automatique. Thèse doctorale, Katholieke Universiteit Leuven.

Monnin, P. and Grosjean, F. (1993). Les structures de performance en français: caractérisation et prédiction. *L'Année Psychologique*. 93. 9-30.

O'Shaughnessy, D. (1981). A study of French vowel and consonant durations. *Journal of Phonetics*. 9. 385-406.

O'Shaughnessy, D. (1984). A multispeaker analysis of durations in read French paragraphs. *Journal of the Acoustical Society of America*. 76. 1664-1672.

Pasdeloup, V. (1988). Analyse temporelle et perceptive de la structuration rythmique d'un énoncé oral. *Travaux de l'Institut de Phonétique d'Aix*. 11. 203-240.

Pasdeloup, V. (1990). *Organisation de l'énoncé en phases temporelles: Analyse d'un corpus de phrases réitérées*, (pp 254 - 258). 18émes Journées d'Etudes sur la Parole. Montréal. 28 - 31 Mai.

Pasdeloup, V. (1992). Durée intersyllabique dans le groupe accentuel en Français. *Actes des 19émes Journées d'Etudes sur la Parole*. Bruxelles. 531-536.

Saint-Bonnet, M. and Boe, J. (1977). Les pauses et les groupes rythmiques: leur durée et disribution en fonction de la vitesse d'élocution. *VIIèmes Journées d'Etude sur la Parole*. Aix en Provence. 337-343.

Thévoz, N. and Enkerli, A. (1994). *Critères de segmentation: Rapport intermédiaire*. LAIP-Lausanne.

Wenk, B. J. and Wiolland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*. 10. 177-193.

Wiolland, F. (1984). Organisation temporelle des structures rythmiques du français parlé. Etude d'un cas. *Rencontres régionales de Linguistique, BLLL* . 293 - 322.

Wunderli, P. (1987). *L'intonation des séquences extraposées en français*. Tübingen: Narr.

Zellner, B. (1994). Pauses and the temporal structure of speech. In E. Keller (Ed.), *Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State-of-the-Art and Future Challenges*. Chichester: John Wiley. 41-62.

ANOTHER TRAVESTY OF REPRESENTATION: PHONOLOGICAL REPRESENTATION AND PHONETIC INTERPRETATION OF ATR HARMONY IN KALENJIN*

John Local and Ken Lodge

Department of Language and Linguistic Science
University of York

1. Introduction

The Kalenjin group of languages, part of the Southern Nilotic or Chari Nile family (Greenberg 1964) are spoken mainly in western Kenya. One of their characteristics is that they display a harmony system which is said to involve the phonological feature Advanced Tongue Root ([ATR]) (Creider and Creider 1989; Hall et al. 1974; Halle and Vergnaud 1981).

In this paper we address issues of the phonological representation of [ATR] in Kalenjin and its phonetic interpretation. Specifically we will show:

- that the harmony system encompasses the C-system as well as the V-system
- that [ATR] is best characterised as a phonological unit which has a syllabic domain
- that there are harmony constraints on the constituents of monomorphemic polysyllables
- that the phonetic exponents of [ATR] harmony provide evidence for the need to maintain a strict demarcation between an abstract, relational phonology and interpretative phonetic exponents (Pierrehumbert 1990; Kelly and Local 1989)

We will argue that one straightforward way of handling the [ATR] harmony system is in terms of underspecification (cf. Lodge 1993b). On

---

* Authors' correspondence addresses: John Local, Department of Language and Linguistic Science, University of York. Ken Lodge, School of Modern Languages and European Studies, UEA, Norwich. NR4 7TJ

York Papers in Linguistics 17 (1996) 77-117
© John Local & Ken Lodge

the assumption that only unpredictable values/features are specified in the lexical entry forms of morphemes (cf. Archangeli 1984, 1988) we will show that

- it is necessary to specify lexically [+ATR] for the dominant morphemes and [-ATR] for the opaque ones.
- the adaptive morphemes are unspecified for lexical [ATR] value.
- [+ATR] harmony domains are immediately adjacent. (There is no evidence that harmony patterns can or do 'skip' over adjacent morphemes.)
- [+ATR] harmony domains encompass immediately adjacent unspecified adaptive morphemes or the default value, [-ATR], applies.

We will propose that a formal implementation of our analysis can be constructed in terms of constraints on structured hierarchies of features which permit partial specification and structure sharing, combined with a phonetic interpretation function (Coleman 1992a; Local 1992; Ogden 1992; see also Bird 1990; Broe 1993; Scobbie 1991).

2. Phonetic interpretation of [ATR]

We begin with a consideration of some of the phonetic characteristics of the [ATR] harmony system in Kalenjin[1]. We will, in the manner of Firthian Prosodic Analysis, refer to these as 'phonetic exponents' (Carnochan 1957; Firth 1948;.Henderson 1949; Sprigg 1957). Importantly our investigations reveal that the phonetic exponents of the [ATR] feature in Kalenjin are varied and not simply confined to the V-system (a detailed discussion is presented in Local and Lodge (forthcoming)). The transcriptions in (1) give an impression of some of these characteristics:

---

[1] The data we discuss is drawn from observations and recordings of a female and male speaker of the Tugen dialect. Both speakers are in their mid 30's.

(1)

+ATR words		**-ATR** words	
[ḳʰȩːβịtʲʰ]	{TO SPRINKLE}[2]	[kʰ·ɛ̧ːβɪṭˠʰ]	{TO GROW}
[ḳʰȩ̣ᵊgụ. tʲʰ]	{TO SCRAPE UP}	[kʰ·ɛ̣ːɣ. ʊ̣ṭˠʰ]	{TO BLOW}
[ḳʰȩːβọḷ]	{TO DIG UP}	[kʰ·ɛ̧ˀβ. a̱. l]	{TO DIG}
[pʰȩ· ɲ]	{MEAT}	[pʰ·ɛɲ]	{HARDSHIP}
[lọ.]	{FAR}	[lˠɔ̣]	{SIX}

2.1 Phonetic differences between words of the [±ATR] categories

There are a number of phonetic differences between words in the two categories which can be observed not only in vocalic portions but also in the consonantal portions of such words. These differences include phonatory quality, vocalic and consonantal quality and articulation and durational differences.

2.1.1 Phonatory differences

The two sets of words exhibit different kinds of phonatory activity. This is audible in terms of voice quality. Words of the [-ATR] set have audible breathy phonation as compared with words in the [+ATR] set. This breathy voice quality is especially noticeable in the rime of the words. Measurements of the open quotient (OQ) of the glottal cycle made from electrolaryngographic recordings (Davies et al. 1986; Howard et al. 1990; Lindsey et al. 1988) and inverse filtering (Karlsson 1988; Wong et al. 1979) show statistically significant differences can be taken to confirm breathiness of phonation (typically, larger OQ values are found for [-ATR] words). Examination of voice source measurements also suggests different kinds of laryngeal behaviour in moving from voice to voicelessness in the two sets of [ATR] words. In [+ATR] voicing dies away slowly and continues at low level (often noticeably overlapping with friction if present). In contrast, in [-ATR] words, voicing drops off rapidly.

---

[2] We adopt the following notational conventions in presenting the Kalenjin material: [phonetic font] for phonetic material; **bold** for phonology; lower case for syntactico-morphological categories; **{bold in braces}** for morphemes expressed in terms of phonology; {CAPITALS IN BRACES} for meanings and glosses. These conventions are based on those employed by Carnochan, 1957. Thanks to Richard Ogden for comments and suggestions concerning notation.

Examination of the spectral characteristics of vocalic portions of the two classes also reveals differences commensurate with breathy versus non-breathy phonation (Local and Lodge, forthcoming). There is, for example, a tendency for words of the [-ATR] set to display a greater amplitude of the fundamental in respect of the first harmonic.

2.1.2 Vocalic differences

There are striking auditory differences in vocalic quality between words in the two sets. Vocalic portions in [-ATR] words are noticeably more central (and frequently more open) than those in [+ATR] words. (Note the open [+ATR] vocoid has a *back* quality in the region of CV5 [ɑ] while the open [-ATR] vocoid has a noticeably *front* quality in the region of CV4 [a]. These harmonize with appropriate tokens from the [ATR] sets: [sɑ̣m^{j}ı̢s^{j}] ~ [ʂa̱.m^{ɣ}ɪʂ^{ɣ}] [t^{h}ɑ̣ŋgus^{j}] ~ [t^{h}·aŋ̄ḡʊʂ̢^{ɣ}].) Examination of plots of F1/F2 for tokens each of the [±ATR] vocoids in the data confirms the results of impressionistic listening (for example, [+ATR] vocoids show lower F1 values than their congeners [-ATR]). For purposes of broad transcription we represent the vowels of Kalenjin thus: [+ATR] [i e ɑ o u], [-ATR] [ɪ ɛ a ɔ ʊ].

2.1.3 Consonantal differences

Words of the two categories exhibit differences in types of consonantal stricture and their ranges of variation. In [+ATR] words we final labial, apical and velar closure with burst release, or with close approximation; in comparable [-ATR] words closure with burst release is not found. In such words lax fricative portions occur but so do portions with open approximation.

There are also noticeable variations in terms of place of articulation. 'Coronals' in [+ATR] words are exponed with apico-alveolar strictures whereas they may be exponed with either apico-alveolar or dental strictures in [-ATR] words. Generally consonantal pieces in [+ATR] words are tenser than their [-ATR] equivalents. This can give rise to the percept of stop-like release of laterals and nasals in [+ATR] words.

2.1.4 Durational differences

Consonantal and vocalic portions are durationally different in [±ATR] words. Typically consonantal portions are shorter in [+ATR] words than in [-ATR] words. This is particularly noticeable in the closure and release phases of initial and final plosive portions. Averages of vocalic duration reveal a tendency for [-ATR] vocoids to be shorter than [+ATR] vocoids but there is some overlap in terms of the ranges of duration. However, [+ATR] words are routinely longer (measured from beginning to end of voicing) than are [-ATR] words.

3. Phonological preliminaries: some characteristics of [ATR] domains

Having provided a brief characterisation of the phonetic exponents of [ATR] we now provide an outline of the main aspects of the organisation of the [ATR] harmony system in Kalenjin. There are three different types of morpheme: *adaptive*, *dominant* and *opaque* whose behaviour can be described as in (2) below:

(2)

(i) dominant morphemes are always [+ATR]; any immediately adjacent adaptive morpheme(s) will share this value: **{MORPH}$_D$**.

(ii) adaptive morphemes vary their [ATR] value according to the specification of [ATR] in their neighbouring morpheme(s): **{MORPH}$_A$**.

(iii) opaque morphemes are always [-ATR] and do not vary the value, even next to a dominant morpheme. They delimit the domain of dominant morphemes: **{MORPH}$_O$**.

3.1 Examples of ATR patterning

In (3) - (8) below we give examples of each of these possibilities with accompanying broad phonetic transcriptions.

(3)

{KE:R}$_{D}$	**{-UN}**$_{A}$	[ke:run]
{SEE} root	directional suffix	{SEE IT FROM HERE}

(4)

{KU:T}$_{A}$	**{-UN}**$_{A}$	[kʊ:tʊn]
{BLOW} root	directional suffix	{BLOW IT HERE} (imperative)

(5)

{KA-}$_{A}$	**{A-}**$_{A}$	**{KU:T}**$_{A}$	**{-E}**$_{D}$	[kɑ:ɣu:te]
recent- past prefix	1sg subject prefix	{BLOW} root	continuous suffix	{I WAS BLOWING}

(6)

{KA-}$_{A}$	**{A-}**$_{A}$	**{KU:T}**$_{A}$	**{-UN}**$_{A}$	[ka:ɣʊ:tʊn]
recent- past prefix	1sg subject prefix	{BLOW} root	directional suffix	{I BLEW IT}

(7)

{KI-}$_{A}$	**{A-}**$_{A}$	**{UN}**$_{D}$	**{-KEJ}**$_{O}$	[kiɑungɛɪ]
far-past prefix	1sg subject prefix	{WASH} root	reflexive suffix	{I WASHED MYSELF}

(8)

{KA-}$_{A}$	**{KA:-}**$_{O}$	**{KO-}**$_{A}$	**{KE:R}**$_{D}$	**{-A}**$_{A}$
recent-past prefix	perfective prefix	aspect prefix	{SEE} root	1sg object suffix

[kaɣa:ɣoɣe:rɑ]

{HE HAD SEEN ME}

Evidence for the three types of morpheme is as follows. Sentences (3) and (4) show that the directional suffix **{-UN}$_{A}$** is an adaptive morpheme; in (3) it appears in [+ATR] form and [-ATR] in (4). Similarly comparison of (4) and (5) show that the verbal root **{KU:T}$_{A}$** may also vary in terms of [±ATR] characteristics and can therefore be treated as adaptive. In (4) we see that any such adaptive morphemes not in the domain of dominant ones exhibit the exponents of [-ATR]. Comparison of the characteristics of the structures in (5) and (6) shows that the continuous suffix **{-E}$_{D}$** is dominant (therefore [+ATR]) and that all the other morphemes in its left domain share its [+ATR] characteristics. In (7) the final suffix is opaque and so it does not share the [ATR] characteristic of the preceding dominant ([+ATR]) root **{UN}$_{D}$**, while the two adaptive prefixes in the left domain of the root share its [+ATR] properties. In (8) the perfective prefix **{KA:-}$_{O}$** is opaque and thus the adaptive recent-past prefix **{KA-}$_{A}$** at the beginning of the construction is outside the domain of the dominant root **{KE:R}$_{D}$**. As expected from the behaviour of the adaptive suffix in (4) this initial prefix is [-ATR]. However, the adaptive morphemes in the immediate left and right domains of the dominant root share its [+ATR] characteristics. Note that roots (nominal and verbal) and affixes may be dominant or adaptive. Affixes may be opaque but roots are not.

[ATR] functions in a variety of ways in Kalenjin. In addition to the harmony patternings in (3) - (8) and the lexical pairs given in (1) above, it participates, for instance, in some singular/plural distinctions: [sọm^{j}i̤s^{j}] {AWFUL} (plural) is [+ATR]; [s̨a̤ m^{ɣ}ɪs̨^{ɣ}] {AWFUL} (singular) is [-ATR]; [mʔ̣ọ:i̤] {CALVES} is [+ATR] ~ [m^{ɣ}ɔ̨:ɪ] {CALF} is [-ATR] (see also Tucker and Bryan 1964).

4. Abstractness of phonological categories: [ATR] and the inadequacy of intrinsic phonetic interpretation

[ATR] harmony is canonically the kind of phonological organisation which has been seen as a candidate for autosegmental status[3] (Clements 1976, 1981; Kaye 1982). We will discuss one such treatment of Kalenjin [ATR] below. However, it is appropriate here to consider

---

[3] Or within the Firthian tradition as 'prosodic'.

briefly one issue which [ATR] harmony in Kalenjin raises for an autosegmental analysis - that of the phonetic implementation or interpretation of the phonological feature [ATR]. While conventional non-linear approaches may be able to characterise graphically the long-domain implications of [ATR], it is not immediately clear how such phonological approaches could deal in any coherent way with the phonetic implementation of an [ATR] autosegment in Kalenjin given the range of different phonetic exponents we have outlined above. The problem arises because in contemporary autosegmental approaches phonological features are deemed to have intrinsic (or intuitive) interpretation — the IPI hypothesis (see eg Clements (on IPI in feature geometry) 1985[4]; Durand 1990; Goldsmith 1990; Pulleyblank 1989). The intrinsic approach to phonetic interpretation represents a continuity of practice from traditional generative phonologies. In the generative tradition phonetic interpretation is merely the end point of a process which maps strings to strings. Phonological representations are constructed from features taking binary values; phonetic representations employ the same features with the difference that they usually take scalar values. In the *locus classicus* of generative phonology, Chomsky and Halle explicitly embrace this view of a phonetics-phonology continuum and write 'We take 'distinctive features' to be the minimal elements of which phonetic, lexical and phonological transcriptions are composed' (1968: 64). This undefended position is only made possible in SPE, as in more recent autosegmental approaches, because there is no attempt at an explicit formulation of phonetic interpretation. In the present case it would require a certain amount of ingenuity to postulate an [ATR] autosegment and find what there is in common between devoicing of coda approximants, breathy voice quality, front or back secondary articulation, consonantal length, particular ranges of consonantal variability and any putative advanced position of the tongue root.

---

[4] Although Clements argues that the geometric organisation of features 'depends upon phonological, rather than physiological criteria' (1985: 240) it would appear that the categories he discusses are deemed to have an intrinsic phonetic interpretation.

4.1 Getting the exponents of [ATR] to 'fall out'

It has been suggested to us (van der Hulst, personal communication) that there might be some kind of phonetic/perceptual relationship even in this case which might serve to rescue a conventional autosegmental treatment of [ATR] in Kalenjin in respect of the IPI hypothesis. The suggested solution would be to propose that [±ATR] is exponed by degrees of vocal tract tension with [-ATR] exponed by a generalised 'lax' articulatory setting and [+ATR] by a 'tense' setting (cf. also the description in Hall et al.. 1974: 244, without reference, and Schachter and Fromkin. 1968, on Akan). This might then allow the consonantal and vocalic features we are concerned with to 'fall out' of the categories set up by the analysis.

However, such an analysis merely sidesteps the issue in replacing 'the feature [ATR]' with some other intrinsically interpreted feature [lax]. In itself this begs the question as to why precisely it should be this combination of phonetic features (not universally 'lax') rather than some other that is implicated in the interpretation of [±ATR] (see also the discussion of cross-language differences in the phonetic interpretation of [ATR] harmony in Lindau and Ladefoged 1986). Moreover, such a proposal would not provide a readily accessible account of the durational characteristics of vowels and consonants or the observed variability in the 'coronal' consonants in the two sets. Nor, as far as we can discern, would it give us any analytic leverage on the counter-intuitive phonetic implementation of the open [+ATR] vowel as [ɑ] and the open [-ATR] vowel as [a].

However, the central problem with postulating universal features like [ATR] is that the phonetic and phonological levels are confounded, phonological categories amount to little more than 'rounded up' phonetics and phonetic detail is constantly being made to fit the phonology (e.g. Lindau on 'r-sounds', 1985). Since the phonetic exponents of the harmony system in Kalenjin do not seem to have been investigated thoroughly until our recent paper (Local & Lodge 1994), it is of particular concern that a number of analyses have chosen [ATR] as the phonological designation of the relationships involved.

4.2 Definitions of `[ATR]`

Harmony systems are of central phonological importance in a large number of languages. They typically involve two sets of phonetic exponents which alternate in some way, though not always in the same way across languages. Let us call these sets A and B; thus far there can be little disagreement. In the case of `[ATR]`, however, a search has been made for a common phonetic parameter for the set of exponents of the phonological category by investigating some, but not all, such languages. This search has been limited from the outset by the unwarranted assumption that the commonality resided solely in vowel phoneme inventories.

Research by Stewart (1967), Lindau (1975, 1978), Ladefoged (1964 (on Igbo), 1971, 1972), Lindau et al. (1973) and Painter (1973) on the `[ATR]` harmony systems in languages of the West African Akan family establishes a connection between the vowel qualities in the two such sets and the position of the tongue root. Lindau et al. (1973) show that advancing of the tongue root may also be used as a mechanism to alter tongue height, as in German and some English speakers, without there being any justification for giving the mechanism phonological status (87)[5]. They thus distinguish between those languages which use tongue root position as the basis of a phonological vowel harmony system and those that use it as an articulatory mechanism for raising the tongue body. Lindau (1978) suggests that the important articulatory effect of advancing or retracting the tongue root in general is to change the shape of the pharyngeal cavity and labels the phenomenon `[expanded]`. This is an elaboration of Ladefoged's (1971, 1972) suggestion that there is a phonological (sic) feature `[wide]` covering three states of the pharynx: wide, as in advanced tongue root articulations, neutral, where the tongue root is in its 'normal' position (which may or may not be the position for `[-ATR]`, depending on the language), and narrow, where the tongue root is retracted. The last state may be the equivalent of `[-ATR]`, but Ladefoged exemplifies it with Arabic [ʕ]. Lindau (1978: 553) also suggests that neutral versus narrow is employed in Arabic to differentiate between non-emphatic and emphatic consonants respectively. This is the only reference to consonants in relation to the position of the tongue root.

---

[5] Kenstowicz (1994: 20,22) provides a clear instance of the unwarranted elevation of tongue root to phonological status in his discussion of vowel symbols.

With the basic groundwork set up in this way it is easy to see how phonologists (who have not necessarily investigated the so-called [ATR] languages directly) find the [ATR] feature attractive as a generic binary label for the two sets A and B. There is apparently a simple intrinsic phonetic interpretation of the phonological phenomenon, a convenient isomorphism: an advanced tongue root produces a wide pharynx, which equates with [+ATR] in the phonology (see, for instance, Hall and Hall 1980 who, in discussing [ATR] harmony in Nez Perce, comment that [+ATR] [ɯ] 'follow(s) naturally if the tongue root is in advanced position when /u/ is articulated' (214)). However, if, as might be expected, a phonological contrast is exponed by a constellation of phonetic exponents, it has been traditionally deemed necessary to have a way of determining the choice of which the (single) exponent should be. For example, in Gimson (1962: 90) we are told that with regard to RP pairs of long and short vowels 'the opposition between the members of the pairs is a complex of quality and quantity', but he decides to take length as the phonologically relevant characteristic (ibid.: 93). In Gimson (1945-49) he demonstrates that for native RP speakers vowel quality and the duration of voicing in the rime are the important cues for vowel 'length'; the criteria used to come to a decision in Gimson (1962) seem to be 'tradition' and a language-teaching expedient (cf. 90-93 for the full discussion). These hardly represent substantive criteria for a motivated *phonological analysis*.

In the context of the present paper we need to be convinced that a single cover term is appropriate for the phenomena under discussion. But even if this position is adopted, it is important that the phonological analysis must at least make reference to the wider phonological and grammatical context of the language concerned, rather than relying on the discovery of some common physical denominator (cf Firth 1948).

5. The abstractness of phonological categories

We will start with a matter that concerns the phonetic interpretation of only the vocalic part of the syllable in Kalenjin: namely, the exponents of the open V's. First of all, it is striking to note that in the investigations of those languages which have an open V distinction in [+ATR] and [-ATR] sets e.g. Akan, (see, for instance, Lindau 1975, 1978, Lindau et al. 1973), little is said about their qualities, the non-open vowels being the focus of attention. The pharyngeal cross-sections for the latter show clear distinctions in the position of the tongue root, but there are no such cross-sections for the low vowels, transcribed in Lindau (1975) as [ə] for [+ATR] and [a] for [-ATR], but in Lindau (1978) as [a] and [ʌ], respectively, without any comment, though on the formant chart (Fig.7, Lindau 1978: 552) [a] appears in a relatively back position near to [ɔ], [ʌ] being omitted. In their transcription of Kalenjin Halle and Vergnaud use [a] and [a], respectively, again without elaboration (unfortunately misinterpreted by Carr 1993a: 260-262, as [a] and [ɑ], respectively)[6]. The important point about the Kalenjin realizations of the two harmonic sets, as far as the low vowels are concerned, is that we find the counter-intuitive occurrence of [ɑ] for the [+ATR] open V and [a] for the [-ATR] open V (cf. the relatively detailed transcriptions given at the beginning of this paper). Careful impressionistic observation and acoustic analysis indicates that the backer of the two vocalics co-occurs with vocalic and consonantal portions which typify [+ATR]. In other words, the expected tongue body position on the front-back axis in relation to the assumed position of the tongue root does not occur. Whatever the facts of Akan, in Kalenjin the tongue body position is clearly not *determined* by the size of the pharynx, so, even if we restricted the phonological domain of the harmony system to the vowels, for the low vowels we would need the contrary interpretation of [±ATR] to their interpretation for the non-low vowels - not a happy conclusion for universals of phonetic implementation.

---

[6] Whether [-ATR] is equivalent to a neutral or retracted tongue root is not a question we concern ourselves with in this paper, but the issue has led to the introduction of another feature [RTR] in the analysis of some languages; see Carr, 1993b and references therein.

As far as consonantal articulations are concerned, the available literature does not provide much in the way of indication of what happens to them when the pharynx is wide (see, for example, Ladefoged 1972, or Lindau 1978). A narrow pharynx, as we have already noted, has been implicated in the production of Arabic emphatic consonants. This is of no help in explaining the consonantal articulations we have observed in Kalenjin, nor in explaining the difference in phonation types. It is Stewart (1967: 199) who assumes a relationship between [+ATR] and breathy voice, for which we find no evidence; on the contrary, in our data breathy voice in the sonorants goes with [-ATR]. (Halle and Stevens (1969) also offer a tentative determinate account of the relationship between tongue-root retraction, larynx lowering and phonatory difference, but the work of Lindau and her associates indicates that such an association is casual rather than causal). Similarly, the lenition phenomena and the length phenomena referred to in §2 above and discussed in detail in Local and Lodge (1994) seem to us to have no obvious connection with pharynx width, any more than the fact that in Kalenjin 'coronality' in [+ATR] words has exclusively alveolar exponents whereas in [-ATR] words it varies between alveolar and dental exponents. The only conclusion we can draw is that [ATR] can have no 'basic intrinsic' phonetic interpretation that will allow us to apply it in any meaningful way to the Kalenjin material under discussion here. Rather the interpretation of the abstract phonological relationship designated [±ATR] must be accounted for in explicit statements of temporal and parametric phonetic exponency (Carnochan 1957; Ogden and Local 1995; Sprigg 1957); we cannot appeal to some kind of free-ride intrinsic phonetic interpretation principle.[7] If we adopt this position, of course, it has considerable ramifications for all aspects of the relationship between phonological categories and their phonetic exponents.

---

[7] Compare the statement of Gazdar et al (1985) concerning similar practices in syntax. 'Unlike much theroetical linguistics, it [the GPSG exposition] lays considerable stress on detailed specifications of the theory and of the descriptions of parts of English grammar ... We do not believe that the working out of such details can be dismissed as 'a matter of execution ... In serious work, one cannot 'assume some version of the X-bar theory' or conjecture that a 'suitable' set of interpretative rules will do something as desired ...' (ix)

Rejection of the IPI hypothesis is, of course, aligned with the position of Firthian Prosodic Analysis wherein phonological representations are entirely relational, encoding no information about temporal or parametric events (Carnochan 1958; Firth 1948; Ogden 1993; Ogden and Local 1993, 1995; Sprigg 1957). Under this view the phonological representations are abstract relational structures and are treated as having no intrinsic phonetic denotation. This is different from the view we highlighted earlier which is propounded in a number of contemporary 'non-segmental' approaches where features in the phonology are deemed to embody a transparent phonetic interpretation - typically cued by the featural *name* (e.g. Browman and Goldstein 1986; 1989; Bird and Klein 1990; Sagey 1986. See also the discussion in Keating 1988).

The position we take does not mean that we see no interesting or 'explanatory' links between phonetic phenomena and phonological structures. Rather our claim is that if we wish to develop a sophisticated understanding of the relationships between the meaning systems of a language and their exponents in speech, being forced to provide an explicit statement of the detailed parametric phonetic exponents of phonological structure is an essential prerequisite. The feature labels for phonological units we employ may be given mnemonic labels (e.g. `[ATR]`), but their relation to the phonic substance need not be simple. Because they are distributed over different parts of the syllabic structure, their interpretation is essentially polysystemic (Firth 1948; Henderson 1949; Carnochan 1957). For example, the interpretation of the contrast given the feature label `[+ATR]` or the label `[+nasal]` at a syllable onset need not necessarily be the same as the interpretation of the contrast given the feature label `[+ATR]` or `[+nasal]` at a rime (see also the comments by Manuel et al. 1992 on the phonetic interpretation of 'alveolarity and plosion' in codas of English words). Moreover, the occurrence of the phonologically contrastive feature `[+nasal]` at some point in the phonological structure may generalize over many more phonetic parameters than those having to do simply with lowering of the soft palate. Similarly the absence of a feature such as `[+voice]` does not necessarily mean that the representation generalizes over tokens where there is no activity involving vocal fold vibration - vocalic, nasal and liquid portions typically have regular vocal fold activity, though the phonological representation to which such portions may be referred does not necessarily involve the *feature* `[+voice]` (cf Ladefoged 1977; Local 1992).

The consequence of this argument is that nothing at all hangs on the *name* of a phonological feature (eg `[ATR]`) provided that the canonical naive view of the relationship between phonological categories and phonetic ones is eschewed. That is provided the semantics of the phonological categories is explicitly and formally stated then it really doesn't matter what they are called. All that the 'naming of parts' achieves is some kind of mnemonic shorthand that can, in the worst cases, lead to analytical infelicities. There are two aspects to specifying the semantics: (i) it is necessary to know how the phonological category(ies) in question relate to other phonological categories - that is provide a semantic statement of their place within the phonological systems and structures and (ii) it is necessary to provide an explicit statement of the phonetic interpretation of the phonological categories - this is crucial because, in Firthian terms, it 'renews the connection' (Firth 1957). For instance, Sprigg (1957:107) writes

> '... it is clear that the phonological symbols are purely formulaic, and in themselves without precise articulatory implications. In order therefore to secure 'renewal of connection' with utterances, it becomes nesessary to cite abstractions at another level of analysis, the Phonetic level: abstractions at the Phonetic level are stated as criteria for setting up the phonological categories concerned, and as exponents of phonological categories and terms.'

We return, therefore, to our initial labels A and B. As cover terms for the categories that enter into the phonological system, they are as good as anything else in that they are abstractions from the data without any phonetic content or implication. It seems to us that this is not dissimilar to a much simpler example that relates to the phonological status of a feature `[alveolar]` or a binary equivalent `[+cor, +ant]`, as a definition of English /t d n/. As is well known, these three putative phonological units are subject to (at least) place of articulation assimilation with a following obstruent or nasal (cf. Gimson 1962, and more recent discussions in Local 1992; Lodge 1984, 1992; Nolan 1992); in other words, their exponents, in this respect vary in terms of articulatory place: bilabial, labiodental, dental, palato-alveolar, palatal and velar, as well as alveolar. The only thing these features have in common is that they are all indeed place specifications. Clearly, in such cases as this the alveolar articulatory place descriptor cannot be equated with the phonological category `[alveolar]`. The proposals made by Local (1992) and Lodge (1981, 1984, 1992) involve non-specification of the place feature for such consonants; in addition, in Local (1992) and Lodge (1992) feature-changing rules are excluded entirely from the grammar, as proposed in §8 below, so by having no lexical specification of a place feature for /t d n/ the necessary level of abstraction is achieved: these particular sounds are not defined as alveolar at all, but as those that have no specific place. (For a proposal that this may be a universal feature of coronals, see Paradis and Prunet 1991.) The appropriate place features are supplied by sharing the following obstruent or nasal in particular structural domains, with alveolarity as the default.

However, the case of Kalenjin is more complicated than this, since the phonetic exponents of the terms of the harmony system cannot easily be subsumed under a general heading such as 'place of articulation'.

Fudge (1967) is an early attempt within the framework of generative phonology to introduce phonological primes with no implicit phonetic content (with a reference to Firthian Prosodic Analysis). He states: 'It is ... dangerous and misleading to say that either articulatory or auditory features ARE the phonological elements, unless they correlate so closely that no facts of language are obscured by treating them as if they were the same' (4, original emphasis). The two reasons he gives to support his claim that facts are obscured if one assumes identity of phonetic and phonological features are the matter of biuniquness (discussed also by Chomsky 1964: 75-95) and morphophonemic patterns, some of which are counter-phonetic. The first of these Fudge exemplifies with tone-sandhi in Mandarin, in which Tone 2 followed by Tone 3, and Tone 3 followed by Tone 3 are both realized as a high rising followed by a low rising pitch (1967: 4-7). (There is evidence that such claims trade on less than compelling phonetic observation - and an innocence about interrelationships between levels of analysis. See, for example, Chuenkongchoo 1956, on Thai and Henderson 1960, on Bwe Karen.) The second is exemplified by the Hungarian vowel system, in which phonetic [ɒ] pairs with phonetic [a:] in a harmony system partly determined by lip-rounding or lack of it; they are phonemicized as /a/ and /a:/, respectively. As Chomsky points out (1964: 74; quoted by Fudge 1967: 10), /a/ is 'functionally unrounded but phonetically rounded.' Fudge sees this as a convenient shorthand, but argues that 'it is surely the task of phonology to make classifications on its own terms, to state explicitly what these phonetic-sounding labels ('Rounded' and 'Unrounded', 'Long' and 'Short', etc.) are a 'shorthand' for' (1967: 10). The Hungarian system also contains a situation parallel to the Mandarin tone-sandhi: [i] and [i:] function phonologically as both front and back, another pair of features involved in harmony relations. He then goes on to show how abstract labels - he uses A, B, 1, 2, a, b, (i), (ii) - can be used to define the phonological relations involved, and then interpreted in four ways, by means of four different sets of rules: articulatory, acoustic, auditory and recognitional. We do not want to go into any further details of Fudge's proposals (which are segmentally based), but would like to note in particular what Fudge considers one serious disadvantage of distinctive feature notation, namely that 'systematic phonemic elements and their systematic phonetic counterparts are treated in terms which are formally indistinguishable, and this often forces us to imply that one systematic phonemic element has been changed into another (Tone 3 HAS BECOME Tone 2 in our [Mandarin] example). This is not only undesirable, but also unnecessary, since we do not require complete biuniqueness in our phonology' (1967: 6). We applaud such cautionary remarks, but we find it extraordinary that after nearly thirty years only a few phonologists have started to pay any attention to them.

4.2 Maintaining strict demarcation: Compositional Phonetic Interpretation

We have argued that the IPI hypothesis for phonological categories is, in the general case, untenable and, in the particular case of [ATR] harmony in Kalenjin, demonstrably inadequate. In the light of this we have suggested that it is not only desirable but necessary to adopt an analysis in which a strict demarcation between the abstract phonological and physical phonetic levels is maintained as in Firthain prosodic analysis. In order to do this, as we indicated, it is necessary to solve the issue of the phonetic interpretation of phonological categories. To accomplish this we adopt the proposal of Coleman and Local (1992) for a compositional phonetic interpretation (CPI) function for partial phonological descriptions. We sketch only the broad outlines of the CPI here. Fuller, more technical descriptions, of the phonological theory and the formal treatment of the CPI function, as formally implemented in the YorkTalk speech generation system, can be found in Coleman 1992a; Local 1992; Ogden 1992). .

In the CPI function adopted here, phonological structures and features are associated with phonetic exponents. The phonological descriptions being interpreted are here taken to be unordered acyclical graph structures with complex attribute-value node labels (cf structures found in GPSG or HPSG). The statement of phonetic exponents in CPI has two formally distinct parts: temporal interpretation and parametric phonetic interpretation. Temporal interpretation establishes timing relationships which hold across constituents of a phonological graph while parametric interpretation instantiates interpreted 'parameter strips' for any given piece of structure (any feature or bundle of features at any particular node in the phonological graph). The resulting 'parameter strips' are sequences of ordered pairs where any pair denotes the value of a particular parameter at a particular (linguistically relevant) time. Thus in the general case:

{(node: partial_phonological_description),(Time_start, Time_2, … Time_end), parameter section}

where the node represents any phonologically relevant contrast domain. (Ladefoged 1980, argues for a similar formulation of the mapping from

97

phonological categories to phonetic parameters.) The time values may be absolute or relative, fixed or proportional. The precise physical domain of the parameter strips (eg articulatory, acoustic, aerodynamic) is not of immediate relevance here.

Under CPI, phonetic interpretation of the phonological descriptions is constrained by the principle of compositionality (Partee 1984) which requires that the 'meaning' of a complex expression is a function of the form and meaning of its parts and the rules whereby the parts are combined. Under the present proposal, the phonological 'meaning' of a syllable equals the 'meaning' of its constituents (for a similar approach see Bach and Wheeler 1981; Wheeler 1981; 1988). The compositional principle is instantiated by requiring any given feature or bundle of features at a given place in the phonological structure to have only one possible phonetic interpretation. So, for instance, in the present case the Kalenjin words (i) [k̥^{hw}o̧ .ˑl̥], 'good planters' and (ii) [k^{hw}ˑ ɔ̣ .lˑ] 'plant!' can be given the following Firthian-like, partial representations (similar representations can be found in Albrow 1975; Carnochan 1960):

$$\text{(i)} \quad {}^{[\text{ATR}+]}(\kappa o \lambda) \qquad\qquad \text{(ii)} \quad {}^{[\text{ATR}-]}(\kappa o \lambda)$$

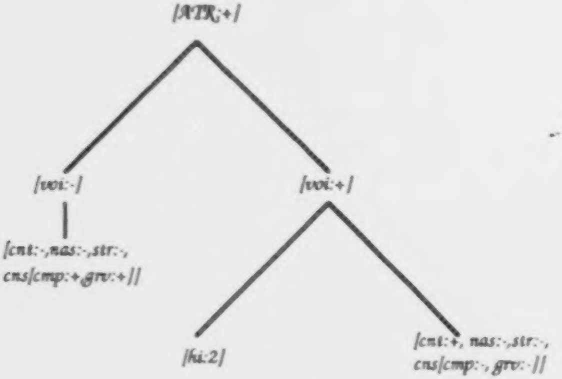
Here the syllable-domain [ATR] unit as well as being semantically distinctive serves to integrate the other syllabic material (paradigmatically contrastive 'phonematic units' (Firth 1948)) with consequences for their phonetic exponency as we illustrated above). Given this, then the interpretation of (i) is of the form:

$$CPI(\,[\text{ATR}:+]\ (\kappa o \lambda)) = \{\text{phonetic exponents of 'kol'}\}$$

where *CPI* is a phonetic interpretation function (cf Coleman and Local 1992). A more fully specified representation of (i) might be given as:

$$\text{(i)} \quad {}^{[\text{ATR}+]}\,({}^{\hbar}(\kappa),\ {}^{\neg\hbar}(o,\lambda))$$

In this representation the units ***within*** the syllable are treated as separate entities or sequences of entities; the superscript symbols $\hbar / \neg\hbar$ placed before the units (κ) and (oλ) serve to indicate onset/rime domain phonation prosodies (*ɦ* 'voicelessness'; ¬*ɦ* 'voice'). Such a representation can be reconstructed as a graph with attribute-value node labels, thus:


[ATR;+]
[voi:-]
[voi:+]
[cnt:-,nas:-,str:-, cns[cmp:+,grv:+]]
[ɦi:2]
[cnt:+, nas:-,str:-, cns[cmp:-, grv:-]]

The compositional interpretation of this schematic representation can be determined in the following quasi-articulatory fashion:[8]

1. *CPI([cnt:-, nas:-, str:-, cns[cmp:+, grv:+]])* = {contact of tongue back with soft palate, closure of soft palate …}
2. *CPI([ɦi:2])*={relatively mid tongue-height…}
3. *CPI([cnt:+, nas:-, str:-, cns[cmp:-, grv:-]])* = {contact of tongue apex with alveolar ridge…}
4. *CPI([voi:+]([ɦi:2], [cnt:+, nas:-, str:-, cns[cmp:-, grv:-]]))* = {succession of *CPI([cnt:+, nas:-, str:-, cns[cmp:-, grv:-]])* to *CPI([ɦi:2])*, relative length of *CPI([ɦi:2])*, relative slow decay of voicing of *CPI ([ɦi:2])*…}
5. *CPI([voi:-]([cnt:-, nas:-, str:-, cns[cmp:+,grv:+]]))* = {voicelessness, aspiration of *CPI([cnt:-, nas:-, str:-, cns[cmp:+,grv:+]])*…}

---

[8] In a more complete representation backness and roundedness of the nucleus would be accounted for at the syllable level, thus providing, *inter alia*, for an appropriate phonetic interpretation of consonant-vowel coarticulation (see Local, 1992).

6. *CPI([atr:+]([voi:-]([cnt:-, nas:-, str:-, cns[cmp:+, grv:+]]), [voi:+]([hi:2], [cnt:+, nas:-, str:-, cns[cmp:-, grv:-]])))* = {succession of *CPI([voi:-]([cnt:-, nas:-, str:-, cns[cmp:+,grv:+]]))* to *CPI([voi:+]([hi:2], [cnt:+, nas:-, str:-, cns[cmp:-, grv:-]]))*, non-maximal backness of *CPI([voi:-]([cnt:-, nas:-, str:-, cns[cmp:+,grv:+]]))* and *CPI([voi:+]([hi:2], [cnt:+, nas:-, str:-, cns[cmp:-, grv:-]]))*, relative palatality of *CPI([cnt:+, nas:-, str:-, cns[cmp:-, grv:-]])*, relative shortness of closure and release of *CPI([voi:-]([cnt:-, nas:-, str:-, cns[cmp:+,grv:+]]))*, tense phonatory quality and slow decay of voicing of *CPI([voi:+]([hi:2], [cnt:+, nas:-, str:-, cns[cmp:-, grv:-]]))*, ...}

We have formally tested and verified a CPI for Kalenjin within the YorkTalk declarative speech generation system employing acoustic parameters. Discussion and illustration of this and quantitative details of the phonetic exponents of [ATR] in Kalenjin are given in Local and Lodge (forthcoming).

6. Phonological analysis

In order to develop our phonological analysis we shall now consider Halle and Vergnaud's (1981) analysis of Kalenjin [ATR] harmony, the contribution of underspecification and then return to a consideration of the phonetic interpretation of [ATR].

6.1. Halle and Vergnaud's analysis

Halle and Vergnaud's (1981) paper was one of the first to argue for an autosegmental account of the Kalenjin harmony system. In it they make a number of substantive claims:

- [ATR] autosegments can be linked only to vowel slots in the core (CV anchor tier), (which they claim is 'obvious').
- [ATR] can also be part of the core specifications, but autosegmental specification overrides core specification.

- Autosegments are either linked to the core in the lexical representations or they are floating, i.e. not linked to the core slots. Linking is subject to the following conditions (= their (1f)):

(9)
i. Each (vowel) slot is linked to at *most* one (harmony) autosegment.
ii. Floating autosegments are linked automatically to all accessible vowel slots.
iii. Unlinked autosegments are deleted at the end of the derivation. (Emphasis original.)

In order to make their analysis work Halle and Vergnaud also find it necessary to invoke the No Crossing Constraint (for a critique of this constraint, see Coleman and Local 1989). To account for the facts in (2) above, as exemplified in (3)-(8), they claim that all vowel slots are (redundantly) specified [-ATR] and that dominant morphemes have a floating [+ATR] autosegmental specification in their lexical entry form. Opaque morphemes are specified with a [-ATR] autosegment. On the basis of this analysis they give the lexical representations in (10a,b,c) (= their (1g); we use Halle and Vergnaud's conventions for representing Kalenjin morphophonology but additionally give broad phonetic transcriptions).

(10a)

```
  kI-a-gɛr          [kɪaɣɛr]               {I SHUT IT}
```

(10b)

```
                 [+ATR]
kI-a-gɛr           -ɛ        [kiɑɣere]     {I WAS SHUTTING IT}
```

(10c)

```
[-ATR]      [+ATR]
   |
ka-ma-a      -gɛːr    -ak  [kamaɑɣeːrak]  {I DIDN'T SEE YOU
                                                  (pl)}
```

In the first case (10a), where all the morphemes are adaptive, Halle and Vergnaud state that the form is 'subject to no modifications and surfaces in its underlying form as far as [ATR] harmony is concerned' (1981: 4), giving [-ATR], the redundant specification of all morphemes. In (10b) all vowels are [+ATR] because (9ii) links the autosegment accordingly. In the third example (10c), which is parallel to (8) above, the last three vowels are linked to [+ATR] by (9ii), but the No Crossing Constraint prevents it from being linked to the first morpheme; given the linking of $\{MA\}_O$ with [-ATR] $\{KA\}_A$ surfaces as [-ATR] (= 'is subject to no modifications').

Since they operate with fully specified underlying forms, the association of the floating [+ATR] autosegment necessarily has the effect of changing the value of the redundant [-ATR] specification of the lexical entry form. It is also the case that the 'blocking effect' of the autosegmental [-ATR] specification of the opaque morphemes is arbitrary, in that in other cases (though not in Halle and Vergnaud's paper) spreading can delink such associations (cf. Broe 1992: 153-154). That is to say, whether spreading can delink or not has to be indicated in a language-specific way, and possibly even a phenomenon-specific way.

Halle and Vergnaud's analysis highlights three problems. The first two are of some generality within conventional autosegmental treatments of languages with [ATR] harmony. First there is an unwarranted assumption that [ATR] associates with vocalic slots only. Second there is a reliance on procedural, feature-changing rules (see, for example, the extensive appeal to 'delinking' and 'deletion' in Goldsmith 1990 and papers cited therein). The third problem concerns Halle and Vergnaud's arbitrary account of the blocking effect of the opaque morphemes. We will deal with the first of these problems in the following section and with the other two when we give a declarative analysis of Kalenjin [ATR] harmony.

7. The syllable domain of [ATR]

It is now appropriate to take a closer look at our earlier claim that [ATR] harmony in Kalenjin is of syllabic domain. Halle and Vergnaud, in conventional manner, associate [ATR] autosegments with vowels (in this way they define dominant morphemes 'those with [=ATR] (sic)

Given that [ATR] harmony systems are conventionally dealt with under the rubric 'vowel harmony' it may seem somewhat bizarre to suggest that there is anything odd about this analytic claim. However, as we indicated at the outset of this paper, the phonetic characteristics of consonantal portions in Kalenjin also show marked differences depending on their occurrence in [±ATR] domains. For example, initial voicelessness and plosion have short voice onset times in [+ATR] domains, but relatively long voice onset times with relatively greater amplitude of burst in [-ATR] domains. In [+ATR] words such as [porpor] ({CRUMBLY}, plural) the apical portion is typically a palatalized trill; in contrast in the [-ATR] form [pɔrpɔr] ({CRUMBLY}, singular), we typically find a velarized tap or a lax apical approximant.

That consonantal portions should be implicated in the exponency of 'vowel harmony' should not be regarded as odd. There is evidence that in other 'vowel harmony' languages consonantal portions may also be different. For example, Kelly and Local (1989: 180) show that in Igbo comparable intervocalic consonant portions vary in a number of ways (e.g. in degree of stricture) according to the harmonic V-system they occur with; Waterson (1956) similarly demonstrates that consonantal portions in Turkish exhibit harmonic properties which go around with the so-called vowel harmony in that language. (Dick Hayward (personal communication) confirms noticeable consonantal differences, particularly in duration, co-incident with the vowel harmony systems in Dinka.)

It is important to stress here that the phonetic characteristics of consonants which we have described are not to be attributed to low-level 'co-articulatory effects' (as might, for instance, be argued in the case of 'emphatic consonant harmony' in Arabic (van der Hulst and Smith 1982)[9]. We therefore contest Halle and Vergnaud's assumption about [±ATR] association. It arises simply because the authors have paid insufficient attention to the phonetic facts of the language.[10]

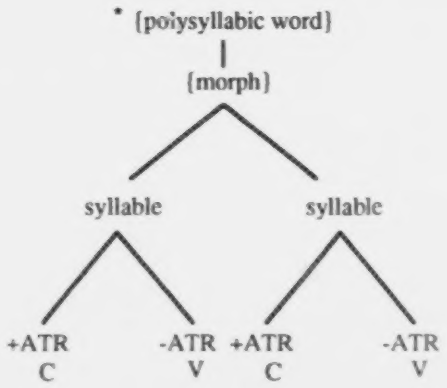
---

[9] Given Whalen's (1990) disscussion concerning the 'planned' nature of so-called low-level 'phonetic coarticulation effects' it is probably dangerous to propose such an account in any case.

[10] This may be a problem of some generality - wherein particular analytic concerns or 'hunches' focus, in an unwarranted and potentially damaging

The situation we have described for Kalenjin is one in which it would be arbitrary to assign the harmony feature [±ATR] to either vowels or consonants. We note, for example, that structural configurations of the kind in (11) are not permitted:

(11)



That is, we do not find cross-combinations of these [+ATR] consonantal portions with [-ATR] vocalic portions or vice versa. We refer to this cohesiveness of [ATR] within syllables as the Syllable Integrity Constraint.

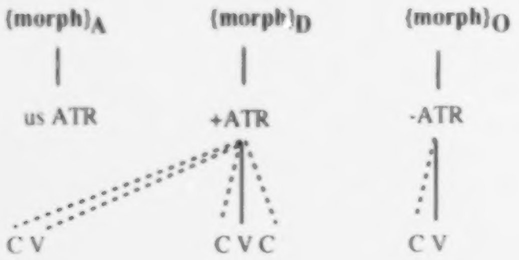
Second, we note here that there are syntagmatic dependencies between onset and rimal constituents and within the rime between nucleus and coda constituents. That is, while we find V, CV, VC as autonomously occurring structures we do not find C (without the implication of a following or preceding V). Taken along with our observations about the integrity of [ATR] in CV(C) structures this suggests that we need to formulate a constraint on the *syllabic* association of [±ATR].

---

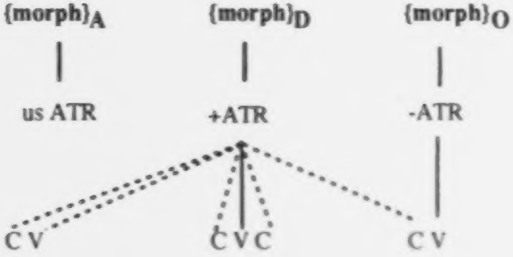
manner, phonetic observation (cf Kelly and Local, 1989). This problem is compounded by the willingness of many current phonologists to 're-work' the *analyses* of others.

We have just proposed that the simplest analysis for the phenomena we have described would be to propose the syllable as the minimal domain of association for [ATR]. We now consider some of the implications of this claim for autosegmental accounts. A conventional non-linear analysis would, like Halle and Vergnaud's, propose association of the [ATR] feature with V-slots and then to allow spreading (cf. also Archangeli 1985; Clements and Sezer 1982; Goldsmith 1990, for example). Notice, though, that we need to deal with two kinds of spreading. While both [+ATR] and [-ATR] spread to all material within syllables only [+ATR] spreads between syllables. Given the inclusion of consonantal material in the 'harmonic spreading' and the Syllable Integrity Constraint, if we adopt the conventional V-association approach, it is clear that we need to invoke a more complex architecture of association precedence and/or blocking to ensure that spreading works in the appropriate fashion. For instance we desire 12(a) but not 12(b).

(12a)



(12b)



In 12(a) we have appropriate spreading of [+ATR] to the C's in the dominant morpheme and to the V and C in the adaptive morpheme (usATR = unspecified [ATR]). This is in line with our observations that it is necessary to spread [±ATR] to any onset and coda consonants as well as vowels, and that dominant [+ATR] harmony spreads to all adaptive morphemes in its domain.

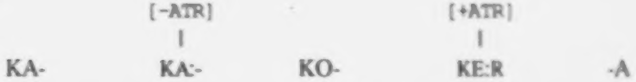
In 12(b), however, although we have spreading of [+ATR] as in (a) to the C's in the dominant morpheme and to the C and V in the adaptive morpheme, it also spreads to the C in the [-ATR] opaque morpheme in violation of the Syllable Integrity Constraint. Clearly we need a way of blocking the spread of dominant [+ATR] harmony to the C's of adjacent opaque [-ATR] syllables. It would be possible to propose a function which would allow morphemic information to percolate to the C and V material in such syllables. However, there is a simpler way of prohibiting this association by ordering the spreading of [±ATR] to C's *within* syllables before spreading between syllables. Once the parochial within-syllable spreading had been accomplished, between syllable spreading would ensure that [+ATR] only associated with V slots which were unspecified for [ATR] and in its immediate left or right domain. This, of course, is tantamount to associating [±ATR] with complete syllables in the first place. As we will show now, it is possible to avoid these somewhat baroque extrinsically ordered association rules if we treat [ATR] as having a syllabic domain and adopt a constraint-based feature-sharing analysis of the harmony system.

8. A declarative underspecification analysis of [ATR] in Kalenjin

One way of avoiding destructive phonological rules, in which features or values are changed or deleted from lexical or, in a derivational framework, intermediate representations, whilst maintaining a single lexical representation for each morpheme, is to employ underspecified lexical representations. Radical underspecification has been developed by Archangeli (1984 1988) and applied to the [ATR] harmony system in Yoruba by Pulleyblank (1988) and Archangeli and Pulleyblank (1989). The Yoruba system that they describe is different in several respects from that of Kalenjin, but the same principles of analysis apply in each case. (In Yoruba, for instance, the vowel /i/ is opaque to the harmony system, whereas in Kalenjin certain *morphemes* are opaque.)

In general, in those cases where alternant realizations are involved, the appropriate feature(s) or feature-value(s) must be unspecified lexically (cf. Lodge 1992 and 1993a). (Whether one refers to features or values is to some extent a matter of whether one uses unary or binary features, respectively; see also the discussion in Calder and Bird 1991. Under these assumptions, then, in Kalenjin the adaptive morphemes are appropriately represented without a lexically specified value for the [ATR] feature underlyingly. Dominant morphemes are specified as [+ATR] (let us say, for the time being, associated with their syllable head (vowel) slot(s), i.e. not floating as in Halle and Vergnaud's analysis). [+ATR], being the non-default value, will have in its domain any adjacent syllables whose head features are not specified for [ATR], i.e. those of the adaptive morphemes. In those words that involve no dominant morphemes, as in (4) and (6) above, a language-specific default rule will supply the redundant specification [-ATR]. (Which value of [ATR] might be the universal default is unclear; in Yoruba, for instance, [+ATR] is the redundant value, though the rule is described as a language-specific complement rule by Pulleyblank 1988: 238, and Archangeli and Pulleyblank 1989: 180, footnote 11.) The opaque morphemes are lexically specified as [-ATR], as in Halle and Vergnaud's account, but given that we have ruled out destructive rules a priori as a means of restricting phonological theory, such lexical specifications will automatically serve to 'block' the 'spread' of any feature, since delinking of any kind is not permitted. Thus, in an underspecification account opaque morphemes are lexically specified for [ATR], whereas adaptive ones are not. This will yield lexical representations of the kind given in (13) for example (8).

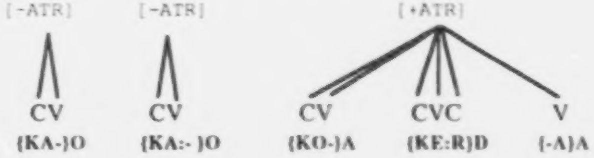
(13)



The unspecified {KO-} and {-A} are in the domain of {KE:R}$_D$ and share its [+ATR] specification. The initial {KA-}$_A$ has the default value [-ATR]. As we demonstrated earlier, this is because the presence of [-ATR] in the lexical representation of the second prefix delimits the inheritance domain of [+ATR].

Since, in the case of Kalenjin, we are dealing with constellations of interacting phonetic parameters which also affect consonantal quality, our analyis above is equivalent to extending the Ladefoged/Lindau proposal to any appropriate consonants, as they do for Arabic. The result is that in Kalenjin the whole syllable is [±ATR] covering both consonants and vowels; our representation in (13) would then be easily modified as in (14), as a representation of the results of spreading and default specification.

(14)



(We do not concern ourselves here with the difference between long and short vowels here, labelling both as V.)

7.1 Structure-sharing,and [ATR] harmony.

In §4.2 we proposed a Compositional Phonetic Interpretation function to allow us a formal means of relating abstract phonological categories to their phonetic exponents. Here we outline a declarative structure-sharing account for [ATR] harmony which is consonant with this CPI.

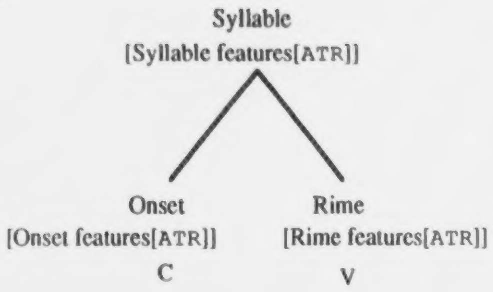
The syntagmatic dependencies outlined above in §7 above imply that V is the head of the syllable rime and that the rime is the head of the whole syllabic structure. This provides us with an obvious solution to the formulation of syllabic association of [±ATR]. In recognising V-system units as *heads* of rimes, rimes as heads of syllables and C-system units as dependents we are able to employ a version of the familiar feature sharing constraints of the GPSG framework (Gazdar et al. 1985). By designating a daughter of a particular category to be the head we identify the relationship between that daughter and the mother as a distinguished one. This allows us to encode the apparent 'feature-spreading' of [±ATR] within a CV(C) structure as a declarative feature-agreement constraint. What we require is to be able to say: OnsetFeatures[ATR] = RimeFeatures[ATR] (and NucleusFeatures[ATR] = CodaFeatures[ATR]). This can be accomplished by employing versions of Gazdar et al's Head Feature Convention (HFC) and Foot Feature Principle (FFP) (Gazdar et al. 1985: 50ff; 70ff). These two constraints may be phrased informally thus for a given fragment of graph representation:

- HFC: The head features of the mother must be an extension of the head features of the head daughter.

- FFP: The foot features of the mother must be identical to the foot features of every daughter.

Combining the HFC and FFP with the structure in (15) below constrains [SyllableFeatures[ATR]] and [OnsetFeatures[ATR]] to be identical.

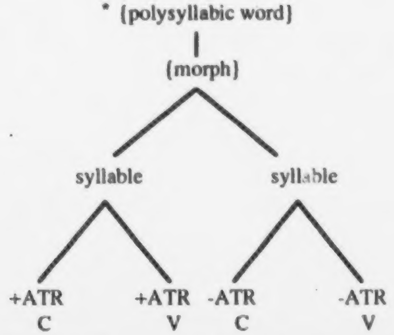
(15)



There are two things to notice here. First observe that it does not matter which of the nodes has its [ATR] value determined or when. The effect is identical (cf Coleman 1992b). Second, notice that the 'spreading' of dominant [+ATR] harmony to immediately adjacent syllables can, by extension, be handled by a similar feature-agreement technique in which the domain of sharing is the word. In Kalenjin a 'word' consists of a monomorphemic root monosyllable or polysyllable. These roots include nominal, verbal, temporal-demonstrative and possessive morphemes (see Lodge 1993b). Roots combine with other morphemes (prefixes and suffixes of various kinds) to form larger word-pieces and these provide the domain of application for the harmony.

Evidence for a word-domain harmony can be illustrated by considering the constraint on the mixing of [+ATR] and [-ATR] vocalic and consonantal portions in monomorphemic polysyllabic structures. Although it is possible, as we have seen in (3) - (8) above, to have polysyllabic utterances in which [+ATR] and [-ATR] properties may be mixed, this is prohibited just in the case where the polysyllabic structure is monomorphemic. So, for instance we find [tɑri:t] {BIRDS} and [tarɪ:t] {BIRDS} where the structures as a whole exhibit [+ATR] or [-ATR] harmonic characteristics. Structures of the following kind are prohibited:

(16)



The ill-formedness of such structure is a natural consequence of the contraint-based analysis we have proposed. Though the syllables respect the Syllabic Integrity Constraint the HFC cannot be satisfied for the {morph} node.

Lodge (1993b) provides further evidence of [ATR] harmony encompassing word-domains. He shows that apparent failures of [+ATR] harmony in some pieces can be attributed to the presence of a word boundary within the piece. For instance, in [kwɛsɑ:ɣɑɲɑ:] in (17), where the syllables are (elsewhere) demonstrably adaptive, dominant, adaptive, dominant, the first syllable would be expected to exhibit [+ATR] harmony features; it does not.

(17a)

$\{\text{KWES}\}_A$	##	$\{\text{NA:}\}_D$	$\{\text{KA}\}_A$	$\{\text{-NYA:}\}_D$
{GOAT} root		temporal demonstrative	recent-past	possessive suffix

[kwɛsɑ:ɣɑɲɑ:][11]

{OUR GOAT (OF
YESTERDAY) }

---

[11] Most sequences of two consonants are not allowed, hence the interpretation of {KWES}+{NA:} as [kwɛsɑ:].

(17b)

{TUKA}$_A$	##	{CA:K}$_D$	{-ET}$_A$	[tʊɣatʃɑ:ɣet]
{COW} root		possessive	recent-past suffix	{THOSE COWS OF OURS }

(17c)

{TUKA}$_A$	##	{-CA:}$_D$	{-KAJ}$_O$	{-KA}$_O$	{-CA:K}$_D$
{COW} root		temporal demonstrative	recent-past suffix		possessive suffix

[tʊɣatʃɑ:ɣaɪɣatʃɑ:k]
{THOSE COWS OF OURS
YESTERDAY}

Similarly in 17(b), [tʊɣatʃɑ:ɣet], where the syllables are adaptive, adaptive, dominant, adaptive, we would expect the first two syllables to harmonise with the dominant syllable, whereas only the last, adaptive syllable harmonizes with the dominant [tʃɑ:ɣ]. If these pieces are analysed as consisting of two words (the second coinciding with the start of the temporal demonstrative in two cases and the possessive in the other), we see that this is exactly the point where the harmony ceases to operate. Once this word division is recognized we find that the harmony operates exactly as it does in (3) -(8).

9. Conclusion

Current work in phonological theory is moving away from procedural, rule-ordered analyses to non-procedural, non-derivational analyses in which phonological representations are incrementally constructed. The phonological representations so constructed cannot be destructively modified - there can be no deletion, 'delinking' or feature-changing rules. The information in the phonological representation must be preserved.

In part, this work represents a research effort to elaborate grammars which favour neither production nor recognition and which allow for a felicitous interaction with contemporary declarative theories of syntax. To this extent, the declarative research program in phonology is a direct descendent of Firthian prosodic analysis (Coleman and Local 1992; Broe 1993; Local 1992; Ogden and Local 1995). The underspecification, feature-agreement analysis we have provided of [ATR] harmony in Kalenjin is intentionally undertaken as part of this research program. Taken together with the Compositional Phonetic Interpretation function which we have described, it provides a more felicitous account of the phenomenon than the mechanisms discussed earlier in the paper and the one offered by Halle and Vergnaud. Unlike the Halle and Vergnaud analysis, underspecification with feature-agreement avoids the need to invoke destructive, structure changing rules. Moreover, in constrast to a conventional V-association account with procedural 'spreading', the feature-sharing constraint offers a computationally tractable mechanism of some generality (Bird 1990; Broe 1993; Coleman 1992b; Local 1992; Scobbie 1991) being more constrained and more comprehensive than a standard analysis in not trading on a naive assumption that the harmony is simply vocalic. In addition to proposing a computationally tractable declarative approach to phonological representation we have also described an explicit declarative, compositional approach to phonetic interpretation which provides the 'renewal of connection' (Firth 1948) between the abstract categories of the phonology and their parametric phonetic exponents.

REFERENCES

Albrow, K.H. (1975) Prosodic theory, Hungarian and English. *Festschrift für Norman Denison zum 50 Geburtstag* (Grazer Linguistiche Studien, 2). Graz: University of Graz. Department of General and Applied Linguistics.

Archangeli, D. (1984). *Underspecification in Yawelmani phonology and morphology*. Unpublished Ph.D. dissertation, M.I.T.

Archangeli, D. (1985). Yokuts harmony: evidence for coplanar representations in non-linear phonology. *Linguistic Inquiry* 16. 335-372.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology* 5. 183-207.

Archangeli, D. and Pulleyblank, D. (1989). Yoruba vowel harmony. *Linguistic Inquiry* 20. 173-217.

Bach E. and Wheeler, D.W. (1981). Montague phonology: a first approximation. In W. Chao and D.W. Wheeler (eds). *University of Massachusetts Occasional Papers in Linguistics*. Volume 7. Graduate Linguistics Association, University of Massachusetts. 27-45.

Bird, S. (1990). Constraint-based phonology. Unpublished PhD thesis, University of Edinburgh.

Bird, S. and Klein, E. (1990). Phonological events. *Journal of Linguistics*.26. 33-56.

Broe, M. (1992). An introduction to feature geometry. In Docherty, G. and Ladd, R. (eds.) *Papers in laboratory phonology II*. Cambridge: CUP. 149-165.

Broe, M. (1993). *Specification theory: the treatment of redundancy in generative phonology*. Unpublished PhD thesis, University of Edinburgh.

Browman, C. P. and Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook* 3. 219-252.

Browman, C. P. and Goldstein, L. M. (1989). Articulatory gestures as phonological units. *Phonology* 6. 201-251.

Calder, J and Bird, S. (1991). Defaults in underspecification phonology. In S. Bird (ed). *Declarative Perspectives on Phonology*. University of Edinburgh. 107-125.

Carnochan, J. (1957). Gemination in Hausa. In *Studies in Linguistic Analysis*. Special Volume of the Philological Society. Oxford: Basil Blackwell. 149-181.

Carnochan, J. (1960). Vowel Harmony in Igbo. *African Language Studies*. 155-63. In Palmer 1970. 222-229.

Carr, P. (1993a). *Phonology*. Basingstoke: Macmillan.

Carr, P. (1993b). Tongue root harmony, lowness harmony and privative theory. *Newcastle and Durham Working Papers in Linguistics* 1. 42-73.

BEST COPY AVAILABLE

Chomsky, A.N. (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, N and M. Halle. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Chuenkongchoo, T. (1956). *The prosodic characteristics of certain partiles in spoken Thai*. Unpublished MA thesis, London University.

Clements, G.N. (1976). *Vowels harmony in nonlinear generative phonology: an autosegmental model*. IULC. Bloomington: Indiana.

Clements, G.N. (1981). Akan vowel harmony: a nonlinear analysis. In G.N. Clements (ed) *Harvard Studies in Phonology* Vol. 2. Distributed by IULC.

Clements, G.N. (1985). The geometry of phonological features. *Phonology Yearbook* 2. 225-252.

Clements, G.N. and Sezer, E. (1982). Vowel and consonant disharmony in Turkish. In H. van der Hulst and N.V.Smith (eds) *The structure of phonological representations (Part I)*. Dordrecht: Foris Publications. 213-255.

Coleman, J. (1992a). 'Synthesis-by-rule' without segments or rewrite-rules. In Bailly, G. and Benoit, C. (eds) *Talking machines*. Amsterdam: North-Holland: Elsevier. 43-60.

Coleman, J. C. (1992b). The phonetic interpretation of headed phonological structures containing overlapping constituents. *Phonology* 9. 1-44.

Coleman, J. and Local, J.K. (1992). The 'No Crossing Constraint' in autosegmental phonology. *Linguistics and Philosophy* 14. 295-338.

Creider, C.A. and Crieder, J.T. (1989). *A grammar of Nandi*. Hamburg: Helmut Buske Verlag.

Davies, P., Lindsey, G.A., Fuller, H. and Fourcin, A.J. (1986). Variation of glottal open and closed phases for speakers of English. *Proceedings Institute of Acoustics* 8. 539-546.

Durand, J. (1990). *Generative and Non-Linear Phonology*. London: Longmans.

Firth, J.R. (1948) Sounds and Prosodies. *Transactions of the Philological Society*, 129-152.

Firth, J R (1957). A synopsis of linguistic theory. *Studies in Linguistic Analysis*. Special Volume of the Philological Society, 2nd edition, 1962, 1-32.

Fudge, E.C. (1967). The nature of phonological primes. *Journal of Linguistics* 3. 1-36.

Gazdar, G., Klein, E., Pullum, G., and Sag, I. (1985). *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

Gimson, A.C. (1945-49). Implications of the phonemic/chronemic grouping of English vowels. *Acta Linguistica* v.

Gimson, A.C. (1962). *An introduction to the pronunciation of English*. London: Edward Arnold.

Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Basil Blackwell.

Greenberg, J.H. (1964). *The languages of Africa*. Bloomington: Indiana University.

Hall, B.L., Hall, R.M.R., Pam, M.D., Myers, A., Antell, S.A. and Cherono, G.K. (1974). African vowel harmony from the vantage point of Kalenjin. *Afrika und Übersee* LVII. 241-267.

Hall, B.L. and Hall, R.M.R., (1980). Nez Perce vowel harmony: an africanist explanation and some theoretical questions. In R.M. Vago (ed) *Issues in vowel harmony*. Amsterdam: John Benjamins.

Halle, M, and Stevens, K.N. (1969). On the feature 'Advanced Tongue Root'. *MIT Research Laboratory of Electronics Quarterly Progress Report* 94. 209-215.

Halle, M. and Vergnaud, J-R. (1981). Harmony processes. In Klein, W. and Levelt, W. (eds.) *Crossing the boundaries in linguistics*. Dordrecht: Reidel. 1-22.

Henderson, E. J. A. (1949). Prosodies in Siamese. *Asia major* I, 189-215. (Reprinted in Palmer, 1970. 27-53)

Henderson, E.J.A. (1960). Tone and intonation in Western Bwe Karen. *Burma Research Society Fiftieth Anniversary Publication* 1. 59-69.

Howard, D.M., Lindsey, G.A. and Allen, B. (1990). Toward the quantification of vocal efficiency. *Journal of Voice*, Volume 4, No. 3. 205-212.

Karlsson, I. (1988). Glottal wave form parameters for different speaker types. *Proceedings of SPEECH 88, 7th FASE Symposium*. Edinburgh: Institute of Acoustics. 225-231.

van der Hulst, H. and Smith, N. (eds) (1982). *The structure of phonological representations (Part II)*. Dordrecht: Foris Publications.

Kaye, J.D. (1982). Harmony processes in Vata. In H. van der Hulst and N. Smith (eds). *The structure of phonological representations (Part II)*. Dordrecht: Foris Publications. 385-452.

Keating, P. (1988). The phonology-phonetics interface. In F. Newmeyer (ed). *Cambridge Linguistic Survey*, vol.1: *Linguistic Theory: Foundations*. Cambridge: Cambridge University Press.

Kelly, J. and Local, J.K. (1989). *Doing phonology*. Manchester: Manchester University Press.

Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford: Basil Blackwell.

Krishnamurthy, A.K. (1992). Glottal source estimation using a sum-of-exponentials model. *IEEE Transactions on Signal Processing* Volume 40. No. 3. 682-686.

Ladefoged, P. (1964). *A phonetic study of West African languages*. Cambridge: Cambridge University Press.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (1972). Phonological features and their phonetic correlates. *Journal of the International Phonetic Association* 2. 2-12.

Ladefoged, Peter (1977). The abyss between phonetics and phonology. In *Proceedings of the 13th meeting of the Chicago Linguistic Society*. 225-235.

Ladefoged, P. (1980). What are linguistic sounds made of? *Language*, Vol. 56, 3. 485-502.

Lindau, M.E. (1975). Features for vowels. *UCLA Working Papers in Phonetics* 30.

Lindau, M.E. (1978). Vowel features. *Language* 54. 541-563.

Lindau, M.E. (1985). The story of /r/. In V.A. Fromkin (ed) *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. New York: Academic Press. 157-168.

Lindau, M.E., Jacobson, L. and Ladefoged, P. (1973). The feature advanced tongue root. *UCLA Working Papers in Phonetics* 22. 76-94.

Lindau, M.E. and Ladefoged, P. (1986). Variability of feature specifications. In J.S. Perkell and D.H. Klatt (eds). *Invariance and variability in* *speech processes*. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 464-479.

Lindsey, G.A., Breen, A.P. and Fourcin, A.J. (1988). Glottal closed time as a function of prosody, style and sex in English. *Proceedings of SPEECH 88, 7th FASE Symposium*. Edinburgh: Institute of Acoustics. 1101-1106.

Local, J.K. (1992). Modelling assimilation in non-segmental rule-free synthesis. In Docherty, G. and Ladd, R. (eds.) *Papers in laboratory phonology II*. Cambridge: CUP. 190-223.

Local, J.K. and Lodge, K.R. (1994). [ATR]: Advanced Tongue Root or Another Travesty of Representation? An investigation of Kalenjin. Paper presented at the LAGB Spring Meeting, University of Salford, April 1994.

Local, J.K. and Lodge, K.R. (forthcoming). On the parametric phonetic interpretation of [ATR] in Kalenjin. *York Research Papers in Linguistics*.

Lodge, K.R. (1981). Dependency phonology and English consonants. *Lingua* 54. 19-39.

Lodge, K.R. (1984). *Studies in the phonology of colloquial English*. London: Croom Helm.

Lodge, K.R. (1992). Assimilation, deletion paths and underspecification. *Journal of Linguistics* 28. 13-52.

Lodge, K.R. (1993a). Underspecification, polysystemicity and non-segmental representations in phonology: an analysis of Malay. *Linguistics* 31. 475-519.

Lodge, K.R. (1993b). Kalenjin phonology and morphology: a further exemplification of underspecification and non-destructive phonology. Paper read to the LAGB Autumn meeting, Bangor, September 1993.

Manuel, S.Y., Shattuck-Hufnagel, S., Huffman, M., Stevens, K.N., Carlson, R and Hunnicut, S. (1992). Studies of vowel and consonant reduction. *Proceedings of the International Conference on Speech and Language Processing*. Volume 2. 943-946.

Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In *Papers in laboratory phonology II*. Cambridge: CUP. 261-279.

BEST COPY AVAILABLE

Ogden, R. (1992). Parametric interpretation in YorkTalk. *York Papers in Linguistics* 16. 81-99.

Ogden, R. (1993b). Where is timing? A response to Caroline Smith. Paper presented at LabPhon 4, Oxford. To appear in A. Arvaniti and B. Connell (eds) *Papers in laboratory phonology* 4. Cambridge: CUP.

Ogden, R. and Local, J.K. (1995). Disentangling autosegments from prosodies: a note on the misrepresentation of a tradition in phonology. *Journal of Linguistics* 30. 477-498.

Painter, C. (1973). Cineradiographic data on the feature 'Covered' in Twi vowel harmony. *Phonetica* 28. 97-120.

Palmer, F. R. (ed.) (1970). *Prosodic Analysis* . London: Oxford University Press.

Paradis, C. and Prunet, J-F. (eds.). (1991). The special status of coronals (Vol.2 of *Phonetics and Phonology*). San Diego: Academic Press.

Partee, B.H. (1984). Compositionality. In F. Landman and F. Veltman (eds). *Varieties of Formal Semantics*. Dordrecht: Foris. 281-312.

Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of Phonetics* 18. 375-394.

Pulleyblank, D. (1988). Vocalic underspecification in Yoruba. *Linguistic Inquiry* 19. 233-270.

Pulleyblank, D. (1989). The role of coronal in articulator based features. in *CLS 25*. Papers from the 25th Annual Regional Meeting of the Chicago Linguistic Society. Part One: The general session. (Eds) C Wiltshire, R. Graczyk and B. Music. 379-393.

Sagey, E. (1986). *The Representation of Features and Relation in Non-Linear Phonology*. PhD. thesis, Massachusetts Institute of Technology.

Schachter, P. and Fromkin, V. (1968). A phonology of Akan: Akuapem, Asante and Fante. *UCLA Working Papers in Phonetics* 9.

Scobbie, J.M. (1991). Attribute-value phonology. Unpublished PhD thesis, University of Edinburgh.

Sprigg, K. (1957). Junction in Spoken Burmese. In *Studies in Linguistic Analysis*. Special Volume of the Philological Society. Oxford: Basil Blackwell. 104-138.

Stewart, J.M. (1967). Tongue root position in Akan vowel harmony. *Phonetica* 16. 185-204.

Tucker, A.N. (1964). Kalenjin Phonetics. In D. Abercrombie, D.B. Fry, P.A.D.MacCarthy, N.C. Scott and J.L.M. Trim *In Honour of Daniel Jones*. London: Longmans. 445-470.

Tucker, A.N. and Bryan, M.A. (1964). Noun classification in Kalenjin: Păkot. *African Language Studies*. 3. 137-181.

Waterson, N. (1956). Some aspects of the phonology of the nominal forms of the Turkish word. *Bulletin of the School of Oriental and African Studies* 18. 578-591.

Whalen, D.H. (1990). Coarticulation is largely planned. *Journal of Phonetics* 18. 3-35.

Wheeler, D. (1981). *Aspects of a Categorial Theory of Phonology*. PhD. dissertation. University of Massachusetts, Amherst. Distributed by the Graduate Linguistic Student Association, University of Massachusetts, Amherst.

Wheeler, D. (1988). Consequences of some categorially-motivated phonological assumptions. In R.T. Oehrle et al. (eds). *Categorial Grammars and Natural Language Structures*.

Wong, D.J., Markel, J.D. and Gray, A.H. (1979). Least squares glottal inverse filtering from the acoustic speech wave. *IEEE Transactions of Acoustics, Speech and Signal Processing*. Volume ASSP-27. 350-355.

ON BEING ECHOLALIC: AN ANALYSIS OF THE INTERACTIONAL AND PHONETIC ASPECTS OF AN AUTISTIC'S LANGUAGE*

John Local and Tony Wootton[1]

Department of Language and Linguistic Science
University of York

1. Preface

A case study is presented of an autistic boy aged 11 years. The analysis is based on audio-visual recordings made in both his home and school. The focus of the study is on that subset of immediate echolalia that has been referred to as pure echoing. Using an approach informed by conversation analysis and descriptive phonetics distinctions are drawn between different forms of pure echo. It is argued that one of these forms, what we call 'unusual echoes', has distinctive interactional and phonetic properties which does not have a counterpart in the speech of non-autistic children. These principally consist of a particular segmental and suprasegmental relationship to the prior adult turn, a particular rhythmic timing and a functional opaqueness. This behaviour is set within the context of this child's general communicative behaviour which, in various ways, places a premium on the use of repetition skills. These skills also inform the child's use of repetition in unusual echoes, though here the interactional and phonetic properties of such

---

* This work was made possible by a grant from the Innovation and Research Priming Fund of the University of York. We would like to thank Kevin, his family, and the staff at his school for allowing recordings to be made, and Fiona Weir for conducting the collection and preliminary investigation of the data discussed here. We are grateful to John Kelly and Patrick Griffiths for their comments on earlier versions.

[1] Department of Sociology University of York

York Papers in Linguistics 17 (1996) 119-165.
© John Local & Tony Wootton

repetitions suggest that they display a distinct interactional stance to the questions that precede them.

1.1 Introduction

Echolalia refers to the repetition of words that have been used by another speaker. It is a phenomenon that has come to have special associations with autism, partly because it often makes up a high proportion of the early speech of those autistic children who learn to speak. The words that the child echoes need not be produced in the immediate context in which the echo takes place. For example, while at home the autistic child can sometimes repeat jingles that s/he has heard on the television on some prior occasion, or phrases that have been heard at school. This type of echoing is often referred to as 'delayed echoing'. It contrasts with those cases in which the source of the words being repeated is in the immediate context. Usually, in the research literature, such 'immediate echolalia' is taken to include child repetitions which are modelled on the prior turn of the child's interactional partner, or the prior turn but one.

Within the literature on autism echolalia is generally viewed as a symptom of this condition. Frith, for example, describes it as 'amongst the most characteristic behavioural abnormalities of young autistic children.' (1989:123). Yet, as Frith and others have noted, forms of repetition akin to immediate echolalia also occur in the speech of normal children. This raises the question of whether there are differences between these two populations with respect to either the nature or frequency of echo usage. The work of Prizant and Duchan (1981) suggests that autistic children may be packaging a wider variety of actions within immediate echo formats. When taking account of non-verbal behaviour, segmental and suprasegmental features they claim to show that seven different functional action types can be reliably discriminated within the overall set of immediate echoes. However, work on normal children between the ages of about 2;0-3;0, the ages at which repetition is most rife, also suggests that various actions can be achieved through repetition formats (McTear 1978; Casby 1986; Greenfield and Savage-Rumbaugh 1993). It may still be possible that there are differences between the nature of these action types in the autistic and normal populations, but for several reasons this is less than clear-cut. The most obvious is that different kinds of speech act classifications have been used in studies of normal and autistic populations. In the light of these and other considerations some writers can still claim that there is little difference in the forms of repetition used by normal and autistic children (Rydell and Mirenda, 1991).

In the course of research on autistic echoing further dimensions of variation within echoes have also been identified. Of special importance is the exactness of the repetition, the degree to which the words in the utterance that is the target of the repetition are reproduced. This parameter is of direct relevance to immediate echoes, and in this respect distinctions have been made between three sub-types. First are 'pure echoes', exact repeats of all or some portion of the words used in the prior target turn. Second are 'telegraphic echoes', repeats of words which are not adjacently positioned in the target utterance. Third are 'mitigated echoes', repeats that include some or all words in the target with additional words added. These three subtypes are illustrated below:

a. Speaker A: Where is daddy's hat
Speaker B: Daddy's hat [pure echo]

b. A: Where is daddy's hat
B: Where hat [telegraphic echo]

c. A: Where is daddy's hat
B: Daddy's hat there [mitigated echo]

Within the autistic population it is the prevalence of pure echoes at a certain stage of development that seems to be the clearest potential case of abnormality in the use of repetition. These pure echoes can preserve suprasegmental features of the target utterance as well as segmental ones, thus giving the impression of a speaker who is simply parroting the speech of the other party. Developmentally such pure echoing gives way to more mitigated forms at later ages, and eventually echoing can be virtually eliminated (Roberts,1989).

Although pure echoing is the example *par excellence* of potentially abnormal echoing behaviour it is not possible to be entirely clear about several of its parameters. For example, we do not know whether the autistic child tends to repeat all the words in the target turn or just some of them. And in the latter case, which undoubtedly occurs some of the time, we do not know which words tend to be picked out for repetition. Their functional properties are somewhat clouded by the fact that their analysis in this respect has usually been combined with the analysis of other kinds of echo, notably mitigated echoes. And, above all, there is still the question as to why this repetition behaviour has the special attraction that it does for the autistic child. To say this, though, is to presume that pure echoes have a special status within the repertoire of the autistic as against the normal child. This, however, is by no means clear. And, if it is the case that the use of pure echoes can serve normal communicative functions among autistic children then we need also to detail the distinctive properties of those that appear abnormal in this regard.

In this study, which is a case study of one autistic child, we will focus principally on the child's pure echoes. We have investigated the different ways in which these echoes can participate in the interaction process, and we attempt to discriminate those that appear to serve a recognisable conversational function from others that seem more equivocal in this regard. In particular we identify a sub-set of pure echoes, ones that we call 'unusual', to which no obvious functional description can be attached. We compare this latter set with comparable instances in studies of normal children so as to decide on whether and in what ways this behaviour is different from potentially analogous behaviour found in normal children. And, in general, we try and situate the child's use of pure echoes within the context of his overall interactional skills and predilections. In this way we arrive at certain conclusions regarding how the child comes to use unusual echoes.

2. The child, the data base and methodological approach

The child, who will be called Kevin, is aged 11 years 4 months at the time when the recordings were made. He lives in England and resides at home with his mother, father and younger sister, attending a school for children with special needs each day. In order to gain an empirical estimation of the degree of Kevin's autism The Childhood Autism Rating Scale (CARS) (Schopler, Reichler, Renner 1986; Schopler, Reichler, DeVellis and Daly, 1980) was applied to over 4 hours of audio-visual recordings of Kevin made in various settings (see below). The result of this rating was 50.5. CARS score of 37-60.0 is allocated to the diagnostic category 'Autistic' and given the descriptive label 'Severe Autism' (Schopler et al, 1986: 57).

Audio-visual recordings were made of this child in a number of different settings. One hour 45 minutes of recording took place in the child's home. Relevant equipment, such as a tripod mounted camera, was made available, and instruction given as to its use. All the recordings were made in the absence of any research worker. The 105 minutes of recording are made up of six sections recorded over two days. They include sections in which Kevin is playing with his younger sister, looking at books with his mother, watching TV with relatives, singing songs with his father and just sitting with his mother and father in the context of no special activity. The other setting in which recordings took place was his school where the recordings were orchestrated by our research assistant. Here we have about 2 hours involving Kevin in an open classroom situation, in various kinds of group work with other children and teachers. In addition, three types of one-to-one session were recorded in the school: a) a 10 minute session between Kevin and a teacher which focussed on word recognition and the assembling of word cards into simple sentences; b) a 14 minute session in which Kevin's mother played a board game with him; and c) 43 minutes in which our research assistant engaged in interaction with Kevin in the context of drawing activity and a large doll's house. For reasons that will be later touched on the various one to one sessions both at home and at school were those that yielded most of the speech on which our analysis focuses.

Table 1 gives an overview of the main forms of speech employed by Kevin on our recordings. The main type of speech excluded from this table is delayed echolalia, speech which did not appear to be addressed to other people with some specific communicative intent and which usually consisted of recognisable reworkings of forms of talk that he had heard on some other occasion. This is excluded from the table partly because it would prove difficult to segment this talk into discrete utterances for the purposes of quantification, and partly because its true extent is difficult to capture from our recordings, especially in the open classroom situation. Very roughly, Kevin's delayed echoing would make up at least as much of his talk as does the category 'Other forms of response to vocal initiation' in Table 1. In addition we have excluded from Table 1 such things as singing and words he says to himself as he is sorting word cards into sentences.

Types of child vocalisation	N	(%)
Vocal initiations	9	(5)
Pure echoes	47	(25)
Mitigated echoes	8	(4)
Telegraphic echoes	0	(0)
Other forms of response to vocal initiation by interlocutor	124	(66)

Table 1. Distribution of Kevin's communicative talk aggregated across a variety of settings.

Our definition of 'pure echoes' is stricter than that generally employed in the literature. It is confined to Kevin's turns which consist exclusively of exact segmental repeats of all or some of the words used in the prior target utterance. The Table conveys very well Kevin's low level of dialogic initiation with other people. Apart from his delayed echolalia most of his talk takes the form of replies to questions. This is true of the various echoes in Table 1 as well as the category labelled 'other forms of response to verbal initiation'. In the main he speaks to others only when spoken to.

Psychometric information about Kevin is not available. It is also difficult to make an informed judgement as to his level of language development on the basis of his vocal output, principally because, as is evident from Table 1, his speech production consists mainly of responses to various kinds of question, which on average fall between 1 - 2 words in length (the mode is 1 word). Both mitigated and pure echoes are always responses to questions, as well as the 'other forms of response' speech. The most advanced of his few vocal initiations is Can I have a crisp please, though we have no means of knowing whether he has control over the syntax involved in the production of such sentences. However, his delayed echolalic speech is generally more complex than that contained in Table 1: here, average utterance length appeared to be between 4 - 5 words. Furthermore, in his one-to-one session with his class teacher he is able to construct, with word cards, sentences like 'daddy and mummy play ball' and 'daddy make tea for me'.

Our approach to the analysis of the data extracts that form the core of this paper is one that is principally informed by work in conversation analysis (Levinson, 1983; Wootton, 1989). This approach insists on the examination of linguistic and other communicative behaviour within its local sequential context of production, and seeks inductively to show how the participants, through the details of their behaviour, adopt particular interactional alignments. Such an approach is, therefore, especially concerned with the sequential position that an utterance occupies, the details of that utterance design (and any co-occurring non-verbal behaviour) and the way in which an utterance is treated by the next speaker. Through the evidence that arises from these details we attempt to construct an analysis that is compatible with the implicit understandings of the participants as they go about their interactional business.

The data fragments are given in a modified form of conventional orthography. Where appropriate for analytic purposes, these are supplemented with impressionistic phonetic information. Segmental information is presented in square brackets following orthographic versions (if such are possible), and pitch information is presented syllable by syllable beneath the relevant turn in inter-linear format where the ruler lines are indicative or top and bottom of the speaker's pitch-range. Certain other conventions are adopted from conversation analysis transcription procedures (Atkinson and Heritage, 1984). These comprise the procedures for depicting speech overlap; the use of '=' to signify no gap between speakers or within the speech of a single speaker; where no pitch transcription is given we use '?' to indicate a general rising pitch contour over a turn (all other turns have general falling pitch); the use of double brackets to enclose transcriber comment; the use of colons to mark sound sustension; (hh) to signify audible aspiration within speech and (he) to signal laughter or chuckling. Timings of pauses are given in seconds; (.) indicates a pause of under half a second.

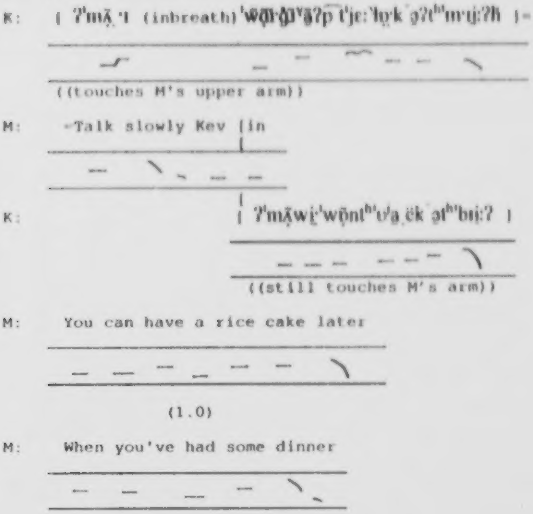
3. General interactional profile

By contrast with normal children the most striking feature about Kevin's verbal behaviour concerns what is absent rather than what is present. Unlike normal children (Snow 1986) he rarely initiates interaction with other people, a pattern that seems as true for his behaviour in his own home as it is for that at school, and a pattern that is characteristic of autistic children more generally (Fay 1988). During free moments at school, for example, he seems content to wander around the classroom, not seeking out contact with other children or staff members, occasionally stopping to look at things, but for the most part absorbed by matters which do not involve direct dealings with other people. His verbal output at such times is made up largely of delayed echolalia; during the recordings this type of talk mainly focuses on regulatory themes. For example, a recurrent utterance frame, both at home as well as at school, is You do not.., articulated with the exaggerated forms of intonation characteristic of an adult reprimanding a child. Typically these utterances are produced on a much higher or lower pitch, and more loudly, than surrounding talk. They exhibit noticeable whispery-voiced phonation and syllable-timing and are often done with dynamic pitch rises on all syllables but the last. Their overall articulatory setting is noticeably tenser than other utterances.

The very infrequent forms of vocal initiation, making up just 5% of his overall vocal output recorded in Table 1, consist exclusively of requests for goods or for the adult to perform an action for him. Sometimes such requests, though still infrequent, can be accomplished in entirely non-verbal ways, as when he takes his mother's hand and moves it towards his back in order to get her to scratch it. When enacted vocally these requests display distinctive articulatory and prosodic characteristics, especially in contrast to the articulatory and prosodic forms that are used to package the remainder of his vocal output. They are produced relatively high in pitch with wide pitch range; any on-syllable pitch movements are likely to be accompanied by noticeable vibrato. The articulatory components are produced laxly and obscurely, the main impressionistic percept being one of overall nasality running through the utterance. These turns also exhibit considerable variations in tempo. Typically they begin slow, accelerate noticeably and slow down. Taken together these phonetic characteristics yield a markedly 'strange' tenor to the speech produced. Kevin's co-interactants orient to the obscurity of utterance and variability of tempo in their talk which responds to these vocal initiations. These features are illustrated in the extract below:

Fragment (1)

Kevin and his mother sit together on the settee at home looking out of the window. His mother looks towards him, but does not speak. Two seconds later he turns to his mother and says, whilst she is still looking at him:



K: [ʔ'mʌ̃ 'ɪ (inbreath) 'wɒ̃ɪ'ŋ̍ɪ'ʌ̃ʔp t'jɛː'lʊ̰k ɡ̥ʔtʰ'mɪŋːʔh]=

((touches M's upper arm))

M: =Talk slowly Kev [in

K: [ʔ'mʌ̃wi̯'wɒ̃ntʰ'ʊ'a̰ ëk ətʰ'bɪiːʔ]

((still touches M's arm))

M: You can have a rice cake later

(1.0)

M: When you've had some dinner

One type of initiation that seems to be entirely absent is that concerned with identifying the names of people or things. Such initiations are commonly enough reported in the literature on normal children, particularly in the kinds of context that frequently occur on our recordings, such as book reading (Ninio and Bruner 1978). In the

BEST COPY AVAILABLE

literature on autism there is some suggestion that the vocal/gestural forms associated with such referential activity are more grossly retarded than, say, those forms associated with the act of requesting (Sigman, Mundy, Sherman and Ungerer 1986; Baron-Cohen 1989). But with respect to pointing, one key ingredient of these referential forms, there is evidence in Kevin's case that he can use this action, together with appropriate vocal accompaniment, to engage in acts of reference. Where he displays this proficiency, however, is in response to questions which seek such a response from him rather than in acts of initiation.

Although the classification of questions that is employed in Table 2 is a fairly crude one it nevertheless suffices to show that the large majority of adult questions to which Kevin gives a non-echoing response are eliciting from him the name of things or persons. Typically these questions take forms like 'What's that?', 'Who is that?', 'Its not a snail its a ?', 'What colour is it?'. For the most part (i.e. 57% of them) they elicit names of things that he can actually see in his surroundings, and such namings are frequently accompanied by points on his part.

Types of information	N	(%)
Visible person/object descriptors	70	(57)
Remote/non visible person/object descriptors	16	(13)
Location descriptions	5	(4)
Course of action information	30	(24)
Other	3	(2)

Table 2. Types of information sought by Kevin's interlocutor in questions which received non-echoing forms of response.

There is ample evidence, therefore, that even though Kevin does not engage in initiating acts of labelling he does, nevertheless, have a wide experience and secure grasp of the labelling game when in response position. In most cases, as in those just discussed in the context of Table 2, when he replies to a question he produces a word that has not been used in the question, he replies in a non-echolalic way. Among the instances of pure echoes, however, there is also evidence of an orientation to and grasp of such a labelling game. Furthermore, the techniques through which such an orientation is displayed suggest that the child has developed quite sophisticated discourse skills in his management of this game.

4. Repetition skills

In this section we will identify various ways in which those who interact with Kevin employ forms of turn design which encourage the use of repetition on his part. In a strict definitional sense his resultant repetitions are often pure echoes, as will be evident from the extracts we use by way of illustration. However, most of these repetitions, by contrast with those we deal with in later sections, appear in no way misfitted for the sequential positions in which they occur, and in most cases they are treated by the child's interlocutor as appropriate moves in the current language game. We begin this discussion by exploring these matters in labelling sequences, ones in which the child is being asked to name something. In assisting the child in his identification of the name in question we shall see that the other party can resort to providing names that the child then goes on to copy.

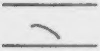
An important general feature of interaction between Kevin and other people is that when they ask him questions he usually does not, initially, give a vocal response. For example, if we take the same questions that form the basis for Table 2, questions that elicited non-echoing forms of response from Kevin, we find that 61% of them occur after at least one prior unsuccessful attempt by his interlocutor to elicit a response to some version of that same question. Indeed, in many cases there are several such prior attempts to elicit a response (e.g. see fragments 3, 5, 8, 9, 11 and 14 below). And this pattern does not seem to be a simple function of the possible difficulty of the question. Questions which seek labels concerning visible objects or persons, perhaps the most straightforward type of question, are preceded by prior unsuccessful elicitation attempts in 60% of cases. If non-response is one type of contingency with which the other party has to deal, a further contingency is that in which the child produces an incorrect response to the question. Most of the questions addressed to him, especially labelling questions, are, of course, test questions, ones for which the other party knows the answer. So the other party can also be placed in the position of guiding Kevin towards the correct answer.

In the context of labelling questions both the contingencies mentioned above, non-response and incorrect response, can be resolved by the other party providing Kevin with a version of the answer that they have been seeking in their question. In fragment (2) his mother says Its jam, while in fragment (3) she says No its a watering can.

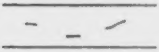
Fragment (2)

Kevin and his mother sitting side by side on the settee at home looking at a book. Kevin begins by correctly identifying a picture of a cake, in response to a question from his mother:

K: Cake

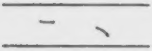


M: A cake with



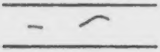
(1.2)

M: What's this ((pointing to, and prodding, a place on the page))

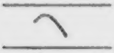


(1.2)

→ M: Its ja:::m=



→ K: =[d̥ʲʑ̥ɛ̥̃·ɜm̥]

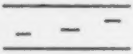


(1.3)

M: So there's ja:m in the ca:ke

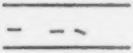
Fragment (3) In same context as fragment (2) above:

M: What is it ((pointing to book))



(1.9)

M: Its a w:: [ɪtsʲə̹wːˑʍˑ]

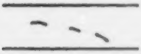


(0.7)

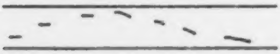
M: • W- [w̃ɔ̰̆ ɦ]

(1.1)

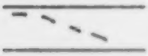
K: [ˈp˭ɛ̨xʐɪk˭öɦ]



→ M: No its a wa:tering ca:n [ˈwˑɔ̣ːtʰə̹ɹˠëŋˈkʰæ̃ːn]



→ K: Watering can [ˈw̤ɔ̤ˑtʰə̹ʊʲɪŋˈkʰˑæ̰̃ŋ̍]

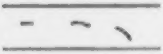


M: What do you do with the watering can?

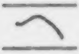
In then producing a repeat of this label in next position, Jam in fragment (2) and Watering can in fragment (3), Kevin is taking this sequential opportunity to produce a first [for him] correct version of the label that the parent has been attempting to elicit from him.In producing this version, then, he is displaying his recognition that this is the appropriate answer. In addition, and as a slight variant of this, Kevin has another way of constructing such repetitions which displays an even closer monitoring of this type of assisting turn.

Fragment (4) In same context as fragment (2) above:

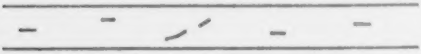
M: What are they ((pointing to book))



K: Berries [ˈbɛ̞ˑʋʲj̞ z̞ʲ]((also points briefly to place on page))

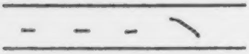


M: They're like berries=they're called



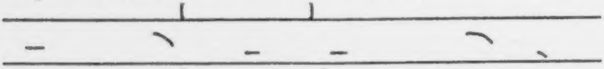
(1.1)

M: What are they called

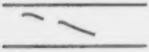


(1.0)

→ M: They're s::tra:[:w b e r r] i es: (.) aren't they

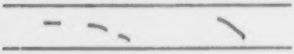


→ K: [Strawb'ries] [ˈlʋ̥ʲɔ̤ː bɔ̣ʋʲɪz̞]
((no point))



(1.6)

M: S:tra:wb'ries (.) Ye::s

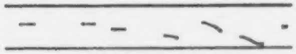


In extracts like fragment (4) he is able to detect from the early part of the word that is produced by the other party, in this case strawberries, what that word is going to be. Indeed, in fragment (4) Kevin also completes the word prior to the completion of the word by his mother. In extracts like (4) the other party can subsequently display some doubt as to the child's grasp of the label in question. In fragment (4) Kevin's mother goes on to say Aren't they (1.6) Strawberries (.) yes, this re-exposure of the child to the correct label perhaps being sensitive to the overlapping position of the child's turn. But in the more frequent cases like fragments (2) and (3) above there is no evidence of these child repetitions being in any way treated as problematic, as displaying some unsound grasp of the language game in question.

A further way in which Kevin can adopt a target word being offered by the adult occurs in circumstances in which the adult offers the child a clue as to the nature of the word being sought. The clue consists of the beginning of the word that the adult is seeking, and such a clue is offered when it has become clear that the child is having difficulty in coming up with the word on his own. In fragment (5), for example, the mother's initial question is answered incorrectly by Kevin, and he is not able to offer an alternative person in response to either of her follow up turns. In this circumstance the mother offers the clue/prompt Aa [ɑ̤̃ːn], which Kevin then manages to complete with tie Sherry [r̃ij̥ˈçëʊʲë] [i.e. 'Auntie Sherry'].

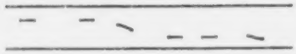
Fragment (5) Mother and Kevin sitting on the settee at home; mother holds a cup in her right hand and has her left arm around Kevin's shoulders, in an affectionate gesture:

M: Who's coming to see you



(1.4)

M: Who's coming to see you ((stroking back of Kevin's neck))

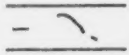


(1.7)

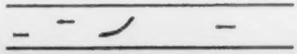
→ M: Aun [ɑ̤̃ːn]

(0.8)

→ K: tie Sherry [r̃ij̥ˈçëʊʲë]



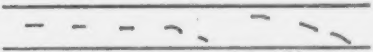
M: Auntie Sherry (.) A::nd?



Similarly, in fragment (6) the child is able to recognise the word that his mother is seeking, 'caterpillar', from her production of the initial voiceless velar plosive of that word. Notice that like the 'Auntie Sherry' instance the child's production of the target word is built as a completion of the prior turn - that is the initial portion is not produced in the child's version.

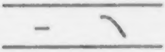
Fragment (6) Mother playing a board game with Kevin in a side room off his classroom at school. Our research assistant is also present. The game involves throwing a dice, which has pictures on its sides. Here his mother encourages Kevin to tell her what the picture is on the exposed side of the dice:

M: Look at the picture what is it=

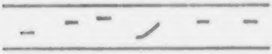


((initially touches his fingers, then points to the dice face in her other hand))

K: =[sɪʔ͡tˑsˈnɛːɹʲə̥ɤ̆] ((briefly points to the dice))



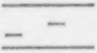
M: Suh not a snail its ak [ɪtsə̥ˈk̥ːʰ]



(1.0)

K: ((obscure quiet)) [kʰə̥ɤ̆]

M: its a [ʔɪtsɔ]?

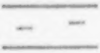


((K briefly points to dice))

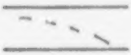
(0.7)

K: Leaf [ˈlə̥j̩ːᶲ]= ((no point))

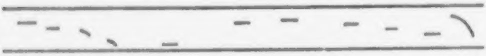
→ M: =it's ak [ĕtsɔʔ·k̥ʰə̥]



→ K: Caterpillar [ˈa̤tʲʰəpʰɪlʲɔ] ((no point))



M: Caterpillar right what have you got to do

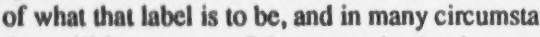


In these various ways, therefore, the child exhibits some skill in monitoring the prior turn of the other party for material that directly cues what is expected of him in his next turn. Routinely, where a label is being elicited the child can look to the prior turn of the parent for a sense of what that label is to be, and in many circumstances, as we have seen, that will be a successful strategy in that it appears to generate a label that is commensurate with the immediate sequential requirements. Labelling games of this kind are important by virtue of their frequency within our corpus of data, but they are not the only ones in which such repetition strategies are fostered. Two further types are now discussed.

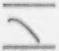
The first is a type of game that is frequently played with Kevin by both his mother and younger sister on our recordings. The game, always initiated by the other party, consists of presenting Kevin with two options and asking him which of these options he would prefer:

Fragment (7) Kevin sitting on the settee at home between his mother and father. Engaged in a playful game in which he is presented with alternatives that he chooses from. The game is already underway when the transcript begins:

M: D'ye wa::n (uh::m) smacked bottom or a kiss?



K: Kiss

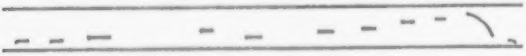


((takes his finger out of his mouth at beginning of this utterance, smiles during it and then angles his cheek to be kissed))

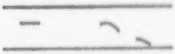
((M kisses K's cheek))

(1.6)

M: D'you wa::nt (.) a smacked bottom or a tickle



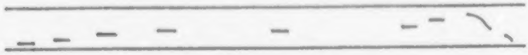
K: Smacked bottom ((smiles during this utterance))



((M playfully smacks his legs, accompanied by
laughter from K and F))

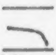
(1.7)

M: Do you wa:nt a: (1.2) ki:ss:: (.) or a tickle



((K's laughter continues through this utterance))

K: Kiss



((turns his head towards M, for kissing, at end of this
word))

Presumably, one feature which makes the game attractive from the point of view of his interactional partner is that it seems to work. It generates serious signs of recognition that Kevin understands the options in question, an understanding displayed partly, perhaps, through his systematic avoidance of certain options, notably being tickled, and through the laughter and horseplay in the course of the game's enactment. Our interest is particularly in the way in which the options are presented. They are both explicitly mentioned by the other party, and characteristically Kevin chooses between the options by repeating the name of that which he prefers. The fact that he does not always select the second of the options with which he is presented is important for later arguments. For now we emphasise that his grasp of the options in question is not just suggested by the considerations above, but also in the minutiae of his non-verbal behaviour: when choosing kiss, for example, his presentation of his cheek for kissing displays an expectation that this will now take place. In these ways his choice of an option is bound up with more than labelling a possibility, it earmarks a course of action that he now expects to take place.

The second interactional tactic with which we will be concerned is also typically used in circumstances in which the other party is seeking guidance from Kevin as to some next course of action. We have already noted that Kevin's co-interactant is often faced with a situation in which no response is made to a question. One course of action that the other party can then use in these circumstances is to transform the question into a yes or no alternative.

Fragment (8) K sitting on settee between his mother and father.

```
     M:    D'you want to go to bed?=

     K:    =[ˈs:ʔ ˈsʲ˔ ˈsʲ˔ [ˈs ˈs ˈs ˈs'] ((then inclines his head more
                           [                                 to M))
     M:                    [Kevin  (.)   Kevi::n

                           (0.7)

     M:    Kevin

                         (1.3)

     M:    Kevin listen (.) [look at me
                            [((puts her hand to K's chin at
                            [  beginning of this
                            [  turn,and directs his face
                            [  towards her))
                            [
     K:                     [ ˈsˠ ˈsˠ ˈs ˠˈsˠ ]

                           (0.7)

     M:    Look at me d'you want to go to be [d
         ((K pushes her hand away from his   [
           chin after word 'me'))            [
                                             [
     K:                                      [ ˈtsʲ ˈsʲ ˈsˠ ]
                                              ((then he looks
                                                away from M))

                           (2.0)
        ((M takes hold of his chin and redirects his face
          towards her))

→    M:   Yes or no

                           (1.1)

→    K:   Yes [ ˈʔjəsˠ ]((as he says this he pulls his chin
                          from her and looks away))

     M:   Ye:s? (.) Are you tired
```

So, in fragment (8), after eliciting nothing other than intermittent voiceless alveolar fricative sounds from Kevin regarding her enquiry as to whether or not he wants to go to bed, his mother eventually formulates the question as Yes or no?. Such a formulation makes it possible for Kevin to answer the original question by picking one or other of the two alternatives, and he responds to this by saying Yes. Here again, then, we find forms of turn design being used by other parties which provide a word that the child can use in coming up with an answer to a question. Indeed, such turn designs might be attractive precisely because they offer such a ready facility to the child.

In his speech with others, therefore, Kevin is mainly concerned with responding to questions, and in the course of this, and in a number of ways, his co-participants offer within their own talk words that Kevin can draw on in constructing a response. In this sense, the availability of repetition to Kevin as a discourse strategy is built into, and fostered, through the turn designs of those he interacts with. And these turn designs are particularly found in circumstances in which the child has not responded or has responded inaccurately. Here, therefore, there is the potential for repetition, as a strategy, to have a particular significance for the child in resolving communication disorder of one kind or another. But its use, as we have seen, is not exclusive to such contexts. In fragment (7), for example, the possibility for repetition to be a viable response is built into the design of turns that are not officially designed to handle a communication problem, and there are other discourse contexts within our data corpus where such is the case. For example, when his teacher asks him to assemble word cards in order to make a sentence she gives him the cards and then vocally models the sentence that he is to make. His job is to reproduce that model, and as he tries to do this he will often say to himself the words that the teacher has used. Here again, as in most of the extracts above, there is little sense of the child's use of repetition being out of kilter with the task in hand. But there are some pure echoes where this is not the case, and it is these which will principally occupy us in subsequent sections.

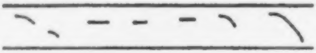
5. Inapposite repetition

In a formal sense many of Kevin's repetitions that we have discussed in the previous section are pure echoes, consisting exclusively of exact segmental repetitions of all or part of a prior adult turn. In the main they appear to be accepted as appropriate conversational moves by the child's co-participant, and in some cases, such as fragment (7) there is good supporting evidence that the child's grasp of the functional role of the repetition is congruent with that of the co-participant. In other cases, however, there might remain doubt as to the kind of understanding displayed through the child's repetition even though the co-participant accepts the child's act as an appropriately fitted conversational move. For example, in fragment (5) it is possible that although the parent is successful in prompting the label 'Auntie Sherry' it may not be the case that Kevin recognises that Auntie Sherry will be coming around later that day. The parent's prompt may simply serve to select one of a number of person descriptors available to the child. And in fragment (8) there is no supporting evidence suggesting that Kevin himself understands that his Yes amounts to an interest in going to bed: for example, on saying this he does not make any physical move which would be consistent with such an understanding.

This kind of semantic/pragmatic insecurity is often tied up with the possibility that at times the child may be operating with a different kind of language game than his recipient. This possibility is concealed, and must remain uncertain, within cases like fragment (5) because the answer that the parent is seeking, 'Auntie Sherry', may also be an answer to an alternative language game that the child might be playing - that of simply guessing which person his mother is referring to. Such a possibility is, however, more clearly realised in other instances like fragment (9) below:

Fragment (9) Kevin sitting on the settee at home between his mother and father. The earlier part of this sequence is transcribed in fragment (13). As the sequence below begins he is sitting with his finger in his mouth, looking frontwards, not at M or F:

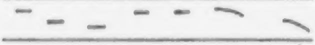
M: Kevin look at my poor cheek



((at the beginning of this turn she touches K's shoulder, then uses that hand to point to her cheek))

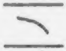
(0.9)   ((K stills his movements here, but does not look at M))

M: Kevin look at my poor che[ek



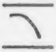
[((initially M touches K's
[hand, which is still in his
[mouth, then points to her
[cheek))

K: [Cheek [tˢɪjʔk]



((turns to look at M, and moves hand from mouth))
K smiles and points at cheek))

M: Look   ((pointing again at her cheek))



Here Kevin's mother is attempting to establish a connection between a mark/stain on Kevin's trousers and some offence that Kevin has committed at an earlier date, an offence which involved his biting her cheek. After initial difficulties in gaining a response from him, and remedial action in the form of touching his hand, Kevin eventually looks at her when she says Look at my poor cheek, words that he can see are also accompanied by a point by her to her own cheek. Kevin's response is to point to her cheek and say Cheek; in fact his production of this word begins prior to his mother's completion of the word Cheek. The fact that he also points to the cheek, that this action is accompanied by a smile and that he just repeats the word 'cheek' (rather than, for example, 'poor cheek') suggests that Kevin's understanding of the sequential expectation obtaining here is for him simply to label the parent's cheek. Just after our transcript ends, once he has become aware of the earlier offence connotations being addressed by his mother and father, his facial demeanour radically changes; pleasure gives way to intense seriousness. And his mother's response to his production of cheek in fragment 9 itself also treats it as misfitted for its sequential position. Her follow up, look, uttered whilst he is already looking at the cheek in question, is clearly attempting to obtain a recognition of the bite related aspect of the cheek.

In this, and other cases, therefore, there is a basis for supposing that the procedure that generates a pure echo on the child's part, the language game that he is playing, can be orderly, though discrepant with that of his co-participant. In fact such discrepancies can appear not just in situations where he produces echoes, they can also be a feature of exchanges in which he produces forms of non-echoing response. For example, in fragment (10) he produces the label Sun in response to his mother's question Listen what have you got to do?, a response that is understandably treated as misfitted to this question by his mother, who reposes it subsequent to his response:

Fragment (10) Mother and Kevin playing the board game at his school: see fragment (6) above for description of the game. Mother is holding the dice, which has a picture of the sun on the top:

```
        M:    Kevin what do you (.) have to do

        K:    ((looks away, then says)) [ s'p̄ɭʊˈlŋʲ ]

        M:    Kevin listen

                        (0.7)

  →     M:    Listen (.) What have you got to do
              ((she taps his hand at word listen, then points to
                top of dice: K's gaze goes to dice))

  →     K:    Sun [ ˈsɔ̨n̆ ]  ((and he points to top of dice))


        M:    You've got to:?
```

Here, as in fragment (9), Kevin's labelling response appears to be cued by the fact that when he turns to monitor his mother's action she is pointing to the focal object in question. His labelling, therefore, arises out of non-verbally influenced understandings of the prior turn of the adult.

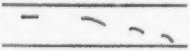
6. Unusual repetition

To this point we have outlined two types of pure echo. In both of these the child's repetition represents a move in a recognisable language game, even though in the second type, just dealt with, such a move is misfitted for the sequential environment in which it takes place. Within Kevin's corpus of pure echoes there remains a further subset that does not fall easily into either of these two categories. This consists of echoes for which a functional description is much more elusive, ones that do not appear to amount to moves in recognisable language games. Indeed, for this reason it may seem somewhat questionable to treat them, as we have done in Table 1, as communicative actions that are commensurate in this respect with the other forms of pure echo. Leaving this issue aside for the moment our initial strategy will be to illustrate this sub-type with two clear examples of it, and then to draw out from these and other examples some general properties of what seem to be these more unusual and puzzling forms of repetition.

The two initial fragments with which we will be concerned in this section are (11) and (12) below:

Fragment (11) Kevin and his mother are in the same board game activity as fragments (6) and (10) above. As this sequence begins M is holding the dice and its container in her hand and K is looking away, towards the camera:

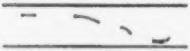
M: Whose turn is it [hʉʐˈtʰɜ̤ːnɪʐˠˈěʔp˺]



((then M adjusts cards on the table between them,
and K looks at the table))

(1.5)

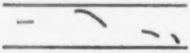
M: Whose turn is it [hʉ̤ ˑʐˈtʰɜːnëzˠˈçʔ͡t˺]



((M manually indicates to table))

(1.5) ((Near end of pause K looks away))

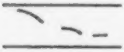
M: Whose turn is it [ˈhʉˑʐˈtʰˈːɜ̤ːnʲɪzʲˈɪʔ͡tʲʰ]



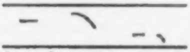
K: ((begins to reach for container M is
holding))

(.)

→ K: Turn is it [ˈtʰɜ̤ːnɪzʲˑɪtʲʰ]((looking at M's face))

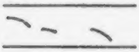


M: Whose turn is it



((withdrawing her hand that holds container))

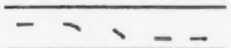
K: Kevin's turn



((his hand now flat on table, not reaching for
container, now looking at table))

Fragment (12) In the same context as fragment (9) above, in fact in the sequence preceding that extract. Kevin has been closely inspecting, and pointing to, one knee of his trousers; as he does this he says quietly, in a tuneful rhythmic way, Doing that doing that on (.) purpose doing that:

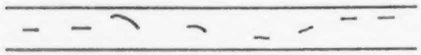
M: Do what on purpose



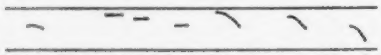
((K then leans back and half looks towards M))

(0.7)

M: Yes you are doing that on purpose

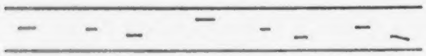


M: you're making a hole aren't you



((as M says this she moves K's hands away from his knee))

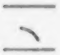
K: Doing a hole doing a hole (in it?)



[ˈɖɔʊ̃ɪŋɐˈɦö̥ˈ dʷu̥ ˜ɪŋɐˈhɔʊ̃ŋ̩ʔ]

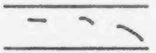
(.)

M: Look ((brief point by M to knee of trousers))



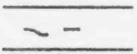
(1.4)

M: Who did that



((sustained point to K's knee))

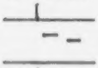
→ K: Who did th- [ˈʒʊ̃ɖɪ̃ˈɦʒ]



((moving his head back sharply))

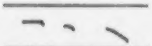
(.)

M: K [evin



[

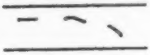
→ K: [Who did that [ˈhɔ̣u̥ ɖɪɖäʔ]



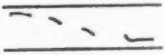
((said as his head comes 'back' to its level position))

(.)

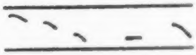
M: Who did it



K: Kevins did it



M: Kevin did it yes



When we speak of these instances as being 'puzzling' we refer in part to the ways in which they are treated by the adult involved. In both these and the other cases in this subset the adult responds to Kevin's pure echoes by reposing the target turn to which the echo was a response. The child's echo is not officially being credited with meaning by the child's co-participant, and in this sense is posing a puzzle to them as well as to the analyst. This way of responding to the child's echoes contrasts with the responses to pure echoes in fragments (2) and (3), that have been previously discussed. But this is only one aspect of their puzzling nature, for we have also seen that some earlier forms of echo are treated similarly by the adult (in fragments (9) and (10)). What makes the echoes in (11) and (12) especially puzzling is that, by contrast with those in (9) and (10), they do not seem to be clear-cut moves in any recognisable language game. This claim needs spelling out a little more, particularly in the light of the analysis of echoes, described earlier, carried out by Prizant and Duchan (1981).

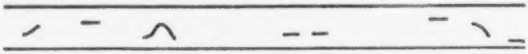
Take fragment (11) above. Here, at the time at which Kevin produces his echo Turn is it, there is direct evidence of a co-occurring hand movement, a right hand reach to take the shaker that is being proffered by his mother, and there is evidence of Kevin orienting to his mother by looking at her. These features should assist in assigning this overall echo configuration into one of the various functional categories outlined by Prizant and Duchan. Yet in various ways this remains a slippery exercise. For example, it could fulfil their criteria for being a request, for being a 'yes' answer to his mother's question or even for being a self-regulatory remark that accompanies his reach. The reaching for the object, for example, could be taken as an affirmation of the fact that he wants it, or it could be taken as evidence of his desire to obtain it. Such matters seem deeply opaque in such instances. Furthermore, it remains a possibility that the child's reach is not strictly connected with the utterance that comes to accompany it. His reach movement begins immediately on the production of his mother's turn, while his Turn is it, together with his gaze switch towards her, is initiated only after she has said the remaining words. So Kevin's overall action configuration could be generated by initially embarking on a course of action, taking the shaker, and then speaking and orienting to his mother on finding himself to be the recipient of her question. In some ways the continuing assuredness of his take attempt and the uncertainty expressed through his continuing gaze at her also speak to such a possibility. Even greater uncertainty features in fragment (12). This time there are no accompanying gestures nor any gaze toward the adult. Kevin's Who did that simply seems to repeat back the adult question, with no obvious indicator of any particular kind of communicative intent.

Therefore, the subset of pure echoes with which we are dealing here has puzzling features both from the point of view of the adult responder and from the point of view of the analyst attempting to engage in functional description. We now turn to describing some typical features of this type of echo.

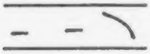
There are three properties of this sub-group of pure echoes which will be addressed. First their segmental correspondence to the model that they are echoing, second their intonational correspondence to this model and third their timing in relation to this model. By segmental correspondence we refer to the fact that the child includes in his echo all the words that occurred in the target/model turn after the initial word that begins the echo. So, in fragment (11) the child could have echoed by saying just 'turn', or by producing a telegraphic version such as 'turn it'. In fact, he produces all the words which occurred in the parental model after his initial word, 'turn' i.e. Turn is it. This is an important feature because we have seen that some of Kevin's echoes can consist of just repeats of non turn final words that are present in the model, notably in fragments (7) and (8). The only exception to this pattern of word inclusion within the present subset of pure echoes is one instance in which Kevin drops an address term that the parent has used in the original model (i.e. the parent says What is it Kevin? and Kevin replies What is it?). From a segmental phonetic point of view, too, these echoes show quite remarkable attentiveness to the articulatory characteristics of the model. Fragment (11) above and (13) below exemplify this close segmental matching. For example, in fragment (11) Kevin's mother's three versions of 'turn is it' are noticeably different in the is it portions. The first is [ɪz̥ˡˠëʔp˺], the second [ëzˠ'ę̈ʔ t˺], the third is [ɪzʲɪ͡ʔ tʲʰ]. The vocalic portions of Kevin's production [ɪzʲɪtʲʰ] have the qualities of his mother's third, rather than first or second version, and the final consonantal portion displays the same front resonance, apicality and aspiration (not noticeable in mother's first two versions) as the immediately preceding version. Similarly, Kevin's echo production of the word boat in fragment (13) shows striking similarities to the preceding adult model rather than to his own prior non-echoed production of the same word:

Fragment (13) Kevin and his mother sitting side by side on the settee at home looking at a book:

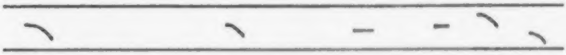
M: oh: what's this (0.1) Kevin (0.1) what is it



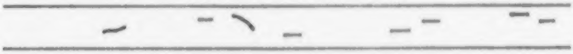
→ K: it's a boat [bö̰ʊ̥ ʔtʲʰ]



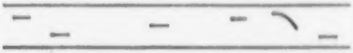
M: boat [bʌ̥ʊtˠʰ] (.) yes (0.2) what's the boat on



M: (0.4) where's the boat on (0.2) Kevin (.) Kevin (.)

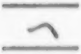


M: oo oo what's the boat on

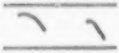


(0.1)

K: river



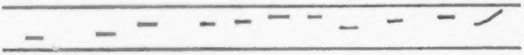
M: river yes



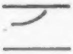
(0.2)'

K: (coughs)

M: would you like to go for a ride in the boat [b̥ʸə̠ʉ̠tʸʰ]

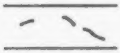


→ K: boat [bʸəʉ̣ʔ͡tʸʰ]



(.)

M: yes or no



We can notice here that Kevin's first production of boat is segmentally different from his mother's in a number of respects. The vocalic portion of Kevin's production has noticeably creaky phonation and begins relatively closer and more rounded than does his mother's; it also finishes noticeably fronter and more open. The syllable coda has co-ordinate glottal closure with the final apical gesture whilst his mother's version does not. The consonantal release of Kevin's production is also noticeably fronter in resonance than that of his mother. Compare this with the phonetics of Kevin's echo which is produced with a vocalic portion and consonantal release which closely match those of his mother's immediately preceding production.

The second property of our subset is the marked tempo, rhythmic and pitch similarity between the echo of the child and that portion of the adult target that is being echoed. Figure 1 below pictures the F0 contours for the relevant parts of fragment (11) (frequency is represented in Hz on the vertical logarithmic axis, time in seconds is represented on the horizontal axis):

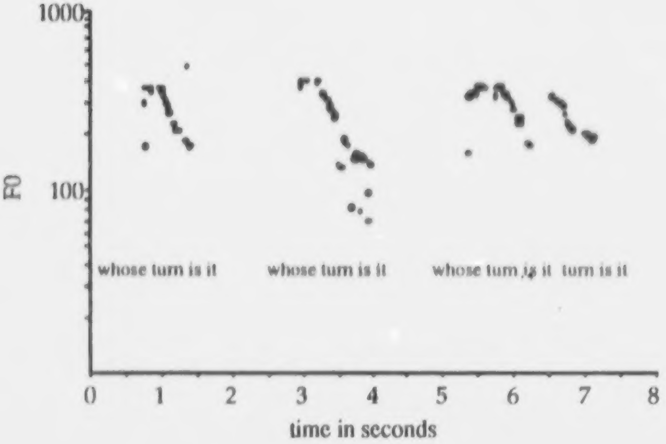


Figure 1 Extracted F0 contours from fragment (11)

We are particularly interested here in the relationship of Kevin's echo 'turn is it' to his mother's third version. There is a close matching of pitch and pitch contour shape (in terms of start and end point; mother's turn is it starts at about 350Hz and falls to around 180Hz; Kevin's begins at about 340Hz and falls to 220H) The durational and rhythmic characteristics of Kevin's turn also model very closely those of his mother's third version. His mother's third version is noticeably slower than the preceding two. The first version has a duration of 835ms with 'turn is it' occupying 572ms The second version has a total duration of 840ms with 'turn is it' occupying 586ms. The third version is 1.22 secs long with the 'turn is it' portion occupying 858ms. Kevin's echoed version of 'turn is it' closely matches this with a duration of 845ms.

Frequency and durational similarities can also be observed in Kevin's repeated version of 'boat' in fragment (13). Extracted F0 contours for the relevant part of this fragment are given in Figure 2 below:

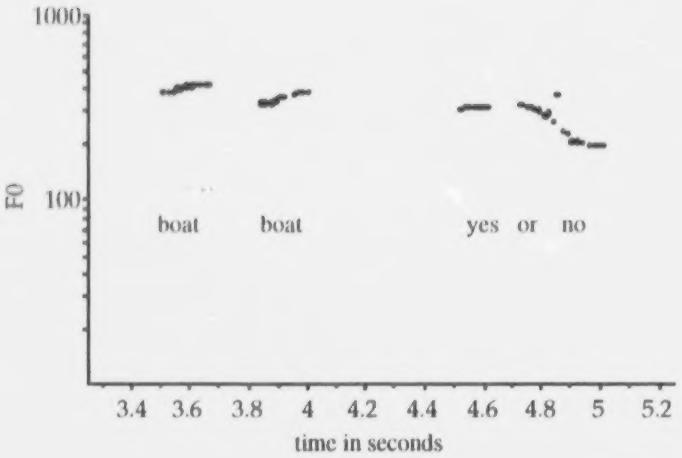


Figure 2 Extracted F0 contours from fragment (13)

Here again there are striking similarities between the pitch configurations of his mother's production of 'boat' and Kevin's version. Both are stepped up rises with initial and final level portions. His mother's production begins at approximately 380 and rises to around 420Hz. Kevin's version starts around 336Hz and terminates around 390Hz. They are also extremely closely matched in terms of their durations: Kevin's lasts 170ms and his mother's lasts 174ms.

In the present data there is at least one instance, in fragment (12), in which the child, on finding his initial echo not being commensurate in these terms with that of the target, redoes the echo so as to produce a version which more closely resembles it. Figure 3 below presents the F0 details for this instance:

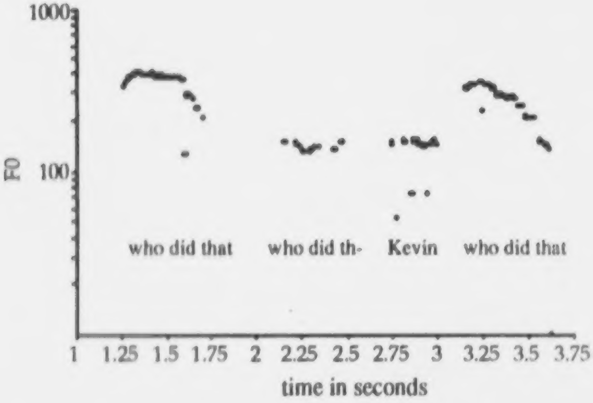


Figure 3 Extracted F0 contours from fragment 12

The child's first production of who did that is done with relatively low level pitch (some 200Hz lower than the starting frequency of his mother's production) which falls slightly towards the end of his utterance (to around 140Hz). It is a quiet, obscurely produced, truncated form of his mother's version. Compare this with his second version which is clearly audible and closely matches the contour and frequency of his mother's version. Mother's version rises from around 330Hz to a peak of 400Hz and falls to around 220Hz. Kevin's second version rises from a starting frequency of around 330Hz to a peak of some 350Hz and falls to about 140Hz. This second version is also more closely matched in terms of duration than his first. His mother's first production lasts some 420ms. Kevin's first version is some 160 ms shorter than this while his second version is 440ms.

It is important to recognise that this phonetic matching is not uniformly found across all instances of repetition produced by Kevin. There are a number of examples where lexically repeated material can be produced with quite different pitch characteristics. The extracted fundamental contours from fragment (2) provide an illustration of this.

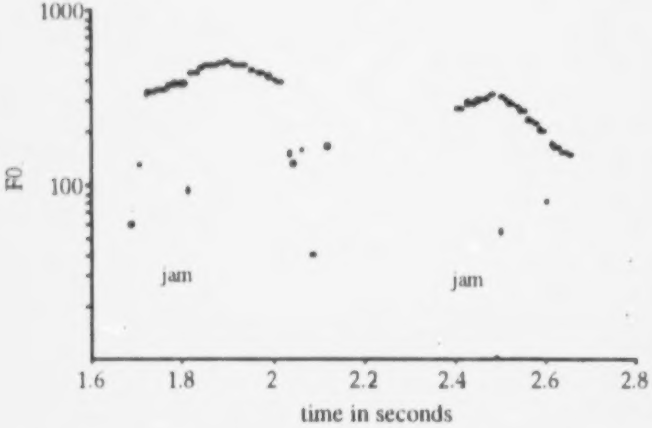


Figure 4 Extracted F0 contours from fragment (2)

Here the mother's and the child's productions are noticeably different. The child's version of 'jam' exhibits a marked fall in frequency towards the end while the mother's does not drop below its starting frequency. The child's version reaches its frequency peak proportionately sooner than the mother's version and shows proportionately less difference in frequency between its starting point and peak. (Mother's version starts around 330Hz rises to 500Hz in some 160ms and falls to 390Hz in 123ms. Kevin's version begins at about 270Hz, rises to its peak of around 320Hz in 57ms and then falls to its end at about 140Hz in 171ms. The amplitude contours of these utterances are different too. In mother's the amplitude peak is skewed towards the middle and end of the utterance. In Kevin's utterance the peak occurs early, closely aligned with the pitch peak, and rapidly falls away thereafter.) The overall duration of the two versions is not matched in the way it is for the 'unusual echoes'. Mother's version lasts 375ms while Kevin's lasts 240ms.

The third feature referred to above concerns those cases where the echo occurs immediately after the adult's target utterance. In this, the normal case, the onset of the echo is routinely rhythmically more rapid than would be expected from the tempo and pattern of rhythm established in the model, a feature which also differentiates this type of echo from several of those discussed earlier in the paper. Couper-Kuhlen (1989, 1990) and Couper-Kuhlen and Auer (1988) provide an innovative and persuasive discussion of such rhythmic organisation in talk. They have shown that turns at talk can be 'contextualised' in terms of their interactional functioning by virtue of their rhythmic constitution and their relationship to the rhythmic patternings in surrounding talk. They demonstrate that if rhythmic isochrony is carefully distinguished from prosodic word stress it is possible to gain an understanding of the kinds of interactional work which can be accomplished by the rhythmic alignment and non-alignment of turns at talk in normal adult speech. This work, based on a substantial amount of natural conversational material, shows that while syllable stress is important for establishing the 'beat of interactional speech rhythm' (1988:4) not all stressed syllables in talk contribute to the perception of rhythmic isochrony. It demonstrates that it is crucially the organisation of talk into isochronous/anisochronous chains, rather than the simple stress patterns of sequences of words which serves to contextualise interactional function. In discussing the rhythmic organisation of question-answer sequences, for instance, Couper-Kuhlen and Auer (1988) observe that:

> 'fillers and vocalisations are not alone indicative of a conversational 'hitch' or, as has been sometimes claimed, of a 'dispreferred' second pair-part. Instead whether or not they are integrated into a larger rhythmic structure seems to affect their conversational function significantly.' (10).

The following two fragments (11') and (13') provide instances of the rhythmic non-integration of the 'unusual repetitions' produced by Kevin.

Fragment (11')

M: Whose /'turn is it
(1.5)
M: Whose /'turn is it
(1.5)
M: Whose /'turn is it
(.)
→ K: 'Turn //is it
M: Whose /'turn is it
K: /'Kevin's turn

The symbol '/' is used to indicate where the rhythmic beat is located; ' indicates prosodic syllable stress.

In Mother's first two turns it so happens that syllable stress and rhythmic beat coincide. In her third turn the rhythmic beat falls in the same place and further reinforces the regular rhythmic pattern established by her first two turns. The stressed syllable 'turn' in Kevin's next utterance, however, is not aligned with this established rhythmic pattern but comes in early. The place where the expected beat would fall is indicated by the symbol '//'. It can be seen that it coincides with the unstressed syllable 'is'. This creates a noticeable anisochronous relationship of Kevin's production with that of his mother's preceding turn. The same phenomenon is evidenced in fragment (13')

Fragment (13')

M: would you /'like to /'go for a /'ride in the
/'boat
→ K: 'boat //
(.)
M: /'yes or no

In this fragment the organisation of Mother's turn is such that the rhythmic beats fall on 'like', 'go', 'ride' and 'boat'. Kevin's turn 'boat', which redoes the final word of his mother's preceding utterance, is not fitted to this rhythmic pattern but again comes in early so that the next beat occurs after the word rather than coincident with its beginning.

When the three kinds of features we have just described combine they give these echoes both a parasitic and autonomic feel. They, like most of the echoes we have been discussing in this paper, are produced in sequential positions in which the child is being required to produce a next turn, but they appear to be occupying that turn simply by repeating a portion of what the adult has said. When these three features are present in the context of single word echoes then, even though the word selected for repetition by the child could amount to an answer to the question, they are routinely treated by the adult as empty and non-meaningful. Nor can the analyst, in such cases, find any basis for supposing that the child has any grasp of the question in hand. Fragments (14) and (15) below illustrate this pattern:

Fragment (14) Kevin sitting on the settee at home, between his mother and father. He has his one arm round his mother's neck; his other hand is holding M's hand throughout the sequence below. His mother has asked him Who do you love?, and Kevin first replies Mummy, then Daddy in response to Who else?. In response to a further Who else? he says Kevin:

M: Kevin ye:(he):s? we know you love Kevin?

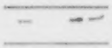
M: (.) Who else

(1.4)

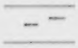
M: What about Lucy

(.)

M: Love Lucy=



• K: =Lucy [lɔʊ'dɪj]



(0.6)

F: Is she asle[ep? (.) Lucy? ((to M))
[
M: [What about Lucy ((to K)) (.) No she's

M: reading ((to F))

F: Oh

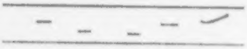
M: What about Lucy ((to K))

(0.8)

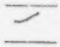
M: D'you love Lucy? ((to K))

Fragment (15) Follows on shortly after fragment (5) above. Kevin and his mother sitting on their settee at home discussing people who might be going to visit them: throughout M is rubbing the back of K's neck with her hand:

```
   M:    And ma:ybe:? (.) Ca: :rl a:s we:ll
  •K:                          ((      ))
                         (.)
   M:    D'you want to see Ca:rl
```



```
•  K:    Carl  ( ŋ'ɒ̤ʱ )
```



```
   M:    Mmmm?=d'you want to play with Carl
                             (0.7)
   K:    ((     ))
   M:    Mm?
```

Although this child is capable of saying 'yes' he does so very infrequently, and some have argued that autistics have special difficulty in engaging in such affirmation (Fay 1988). So, in fragment (14), for example, given this it would be possible for the word 'Lucy' to be an answer to *Love Lucy?*. But presumably the presence of the three features mentioned above in Kevin's *Lucy* leads his mother not to treat his answer as representing his views on this matter: she reposes the question to him by saying *What about Lucy do you love Lucy?*.

There are two further observations that we want to make at this stage about these unusual echoes. The first is that they often do not seem to be associated with questions which are difficult to understand, or ones for which it is difficult to come up with an answer. Notwithstanding experimental work which has shown that autistics are more likely to use echoes after questions that are beyond their understanding (e.g.Paccia and Cursio,1982), there seems nevertheless, in our data, extensive evidence that these unusual echoes are not contingent on the question being ungraspable by the child. This evidence consists of the fact that when the adult reposes the same question to the child after the child's echo then the child often comes up with an answer that is treated as a candidate answer by the adult. In fragment (11) Kevin replies by saying Kevin's turn, and in fragment (12) Kevin's did it. If the question were ungraspable by the child then we might expect to find the child continuing to echo after the adult reposes the question. Importantly, there is one instance of this occurring in our data, so this is a tactic available as a communicative option to the child. But although it is available it only occurs the once. In most cases the child is able to construct an acceptable reply to the reposed question.

Our second, and final, observation in this section concerns the sequential position in which these unusual echoes tend to occur. The observation is that they appear to have a special affinity with the initial stages of any particular line of questioning by the adult. Where they occur they tend to occur as the *first* kind of vocal response that Kevin makes. Logically it would be possible for them to occur in a variety of sequential positions, as do various of those pure echoes discussed in previous sections. For example, after the adult has asked a question and the child has given an initial incorrect response then if the adult reposes the question (e.g. 'No its not an x, what is it?') it would then be possible for the child to produce what we have called an unusual echo, a repeat of the question or some part of it. In practice, however, unusual echoes do not appear in such sequential positions. They are ways of repeating which appear to have their use as a first way of dealing with a question. They are, of course, not the only way of initially dealing with a question. Much more common within these data is non-response on the part of the child. But where they do occur these unusual echoes are usually the first vocal form of response that the child makes to the question.

Before moving on to draw together the various threads of our discussion, with a view to characterising the work achieved through unusual echoes, we first of all want to consider whether it is a distinctive subtype not just in comparison with the earlier types of pure echo that we have discussed but also in comparison with the uses that normal children make of repetition.

7. Repetition in normal children

Within the age range of about 1;6 - 3;0 there is a good deal of repetition within the speech of normal children. Several studies have now shown that turns formatted as repetitions can perform a variety of interactional roles (Casby,1986; McTear,1978). Some of these clearly parallel forms of repetition that we have found in Kevin's data. For example, the use by Kevin of kiss in fragment (7) and Yes in fragment (8) as ways of answering a question follow patterns that are frequent among normal children. The latter can also produce repetitions of what adults say in turns which do not follow overt adult questions. They may choose, for example, to imitate a word that has just been produced by the adult. For example, Casby's (1986) analysis of the talk of one child revealed that 'imitations' made up between 38-49% of all the child's repetitive utterances at MLU stages I-III (using Brown's (1973) criteria for identifying such stages). From the examples of imitation that he provides, like the one reproduced below, it is clear that the child may use the provision of a label by the adult as an occasion for then reproducing this label, either for a first time or with a view to constructing an improved version on their own last try:

Fragment (16) From Casby (1986:136). Mother and child engaged in book reading activity:

```
M:  What's this?
C:  [bʌlaɪ]
M:  Butterfly, right.
C:  Butter-fly
```

This kind of imitative repetition is clearly analogous to forms of repetition that we have found in Kevin's data, notably Jam in fragment (2) and Watering can in fragment (3), and it also informs the more inapposite uses like that of Cheek in fragment (9). Further parallel data among normal children can be found in the more delicate analysis of the language games involved in such situations which is reported in Tarplee (1993). Casby notes (*op cit*:131) that those child utterances he classified as imitative were often intonationally similar to the adult model. This is to be expected in that the child's aim is to produce a version of a word which is similar to that just produced by the adult. Likewise, within our data on Kevin, we have found a tendency for such imitative repetition to be intonationally similar to the target of the repetition (as in fragments (2), (3) and (9)). All in all, therefore, it seems that many of Kevin's pure echoes that we have discussed have their functional counterparts in the language use of young normal children.

What we have described as unusual echoes are answers to questions which do not appear to play a part in any recognisable language game. So, a matter of interest is whether there are counterparts to these echoes in studies of normal children. In order to examine this we will briefly discuss two studies which have examined in some detail particular normal children who have employed repetition as an answering device. Steffensen (1978) describes the answering strategies of two children, one of whom (Jackson), in the age range 1;8 - 2;2 and in the context of yes/no questions, uses repetition rather than 'yes' as a technique of affirmation even though he, like Kevin, is capable of using the negative and affirmative particles. Although such repetitions are often used by Jackson in what Steffensen refers to as semantically well formed ways, ways that are appropriately fitted to the question and which display that the child has some genuine grasp of it, in some cases (such as fragment (17)) this is not so. Steffensen sees such answers as 'responding by formula', as just imitations rather than genuine affirmations, especially when viewed in the light of accompanying nonverbal behaviour:

Fragment (17) From Steffensen (1978:228). Adult and child [Jackson, aged 2;0.7] talk about cutting meat:

```
A:   Shall I cut your meat?
J:   Meat
A:   Shall I cut it?
```

Steffensen's discussion of this child strongly suggests that at a certain stage of development some normal children may resort to using repetition in ways that have some similarities to Kevin's use of unusual echoes. But there are also important actual and possible differences between Jackson and Kevin in this respect. According to Steffensen, a feature of Jackson's repetitions is that they are intonationally different from their models, and in the examples provided by Steffensen there are no cases of the child repeating longer stretches of the question than just a potential answer constituent. Furthermore, there is no discussion of whether, as is the case in Kevin's data (see fragments (11) and (12) above), such repetition answering strategies are also found in response to 'Wh' questions.

A study by McTear (1978) of repetition in his own child between the ages of 2;6-3;1 clearly shows a child who not only produces repetitions of Wh questions but also ones which appear often to include the Wh word itself. An example from McTear is given below:

Fragment (18) From McTear (1978:305): F denotes father, S denotes his daughter who is aged somewhere between 2;6 - 3;1. Presumably, they are talking about what they can see on a television:

```
F: What are they doing?
S: What they doing?
F: They're playing snooker
     ((a few minutes previously S had asked the
     question and received this information))
```

For a variety of reasons, however, these child turns do not seem to us to operate in ways analogous to Kevin's unusual echoes. McTear's argument is that these repetitions are not general answering devices but are specific to particular types of question, what he calls 'display questions'. These are questions in which 'the speaker already knows the answer and wants the hearer to show whether he knows it or not' (*op cit*:302). For McTear the repetition of such questions is a device used by the child to display that she is attending, but one which also intentionally transfers the speaker role back to the questioner. The way that adults are described as replying to these questions supports this contention in that the adult can, after the child's repeat, supply the answer (as in fragment (18)), or the adult can treat the child as deliberately choosing not to answer by insisting on an answer. For example, McTear cites the child's grandmother as responding to such a repeat by saying Come on you tell me (*op cit*:305). Kevin's unusual echoes are never treated in these ways by his co-participant, nor is there ever any clear evidence that for Kevin himself these forms of repetition are designed as speaker switching devices. Furthermore, Kevin's unusual echoes are not specific to particular question types, nor are they, in the main, full repetitions of the prior question. For these various reasons it seems to us that this kind of repetition found in the speech of McTear's daughter is serving a different interactional role than that performed by Kevin's unusual echoes.

8. Discussion

In this article we have been principally concerned with the pure echoes of one autistic boy. Within this relatively unambiguous set of vocalisations we have distinguished three subsets; those which are used in communicatively appropriate ways; those which, though inapposite, represent systematic moves in some language game; and those we have described as 'unusual', that do not amount to moves in any recognisable and conventional language game. We have not quantified these various subsets because their membership is not always clear-cut. For example, our discussion of fragments (5) and (8) has suggested various grounds for uncertainty concerning the kind of understanding that informs Kevin's production of pure echoes in these sequences. Nevertheless, working with what seem to us canonical cases we have tried to identify ways in which these various types are both used by Kevin and responded to by those who interact with him. In doing this we have been especially concerned with the possibly distinctive status of what we have called 'unusual' echoes.

Unusual echoes have a number of features which suggest that they are simply constructed as repetitions of what the adult has said. These features are their segmental and suprasegmental relationship to the model, their unusual rhythmic timing and their functional opaqueness. We have shown, for example, that these unusual echoes appear to be more acoustically matched to their models than is the case for those pure echoes which represent appropriate moves in language games, and that at a segmental level they systematically, and selectively, preserve particular portions of the model. By virtue of these features these unusual echoes impressionistically sound like 'empty' repetition, and are treated as such by the adult. There are, as we have seen in the case of

Steffensen's Jackson, occasional glimpses of somewhat similar behaviour among normal children around the developmental age of about 2;0. But in Kevin's data this type of echo is more intonationally parasitic on the model, not necessarily confined to repeating particular segments of the model and probably more widely used in response to different types of question. As far as we can tell, therefore, unusual echoes do not have counterparts in the speech of normal children.

In developing a characterisation of the role that unusual echoes play in the repertoire of this autistic child it seems to us important to consider them in the context of his more general pattern of interactional skills and involvements. Crucially, vocalisations that are clearly intended as communicative are solicited from Kevin: under 5% of these communicative vocalisations amount-to initiations on his part. His world of spontaneous talk is largely made up of 'delayed echolalia', utterances which are usually recognisable as being authored (Goffman,1979) by other people in other contexts, and ones for which he displays an ongoing, obsessive attachment. It is this domain of language use in which Kevin seems most fluent and at home. And insofar as he rarely displays any continuing and sustained (obsessive) involvement with other people in any particular line of interaction, as evidenced by his gaze, manual behaviour and general bodily orientation, then it seems to be the topics of his delayed echolalia that stand at the forefront of his immediate vocal, and perhaps mental, life.

In these circumstances attempts to elicit responses, communicative speech, from Kevin face the twin tasks of both bringing him out of that separate world and having him understand the import of the adult initiation in question. That the first of these is a problem for those who interact with Kevin is suggested by the frequency with which he appears not to respond to adult initiations, not just in sequences in which echoes occur, but also in those where he eventually makes what is taken to be an appropriate communicative response. The continuing relevance of these considerations routinely occasions various unusual, though for this kind of interaction routine, forms of behaviour on the part of the child's interactional partner - things like emphatic voice, a high frequency in the use of his name as a summons, and physically taking hold of his body so as to encourage orientation to the partner. In the literature more prominence has been given to the second task mentioned above for the adult who attempts to solicit speech from the autistic child, the problem of having the child grasp the linguistic content. Here various research has drawn attention especially to pragmatic and conceptual limitations that make it difficult for the child to understand the nature of what is said to him (Fay, 1988). While this may be so we have argued that this is of limited significance for explaining the occurrence of unusual echoes. The main reason for this is that in many of these sequences, such as fragments (11) and (13), the child seems capable of eventually coming up with an appropriate response to the adult question. Furthermore, it may be important to bear in mind that when asking the child such questions, those who know the child well, such as his mother or a teacher, are unlikely to ask him questions that they know or suspect he is not able to answer, let alone repeat such questions after he produces an unusual echo in response. The key question then, as we see it, is why the child produces such an echo when he has the cognitive equipment to come up with a response?

The answer as to why he chooses to echo seems fairly straightforward. We have seen that the child possesses quite sophisticated skills associated with repetition and that constructing a reply out of material contained in the prior turn is frequently a successful discourse strategy for him in his dealings with other people. And in various ways the design of adult turns, especially in repair sequences after non-response by Kevin, relies on and fosters repetition skills. These points seem to be true not just for the most frequent sequences involving the labelling of things but also in other sequence types such as the games he plays at home. Repetition is thus the obvious device for the child to pick, his most skilled device, in situations which are not conducive to him being able to deal appropriately with an adult question, the situations that seem characteristic of 'unusual' repetition. Much more difficult to specify are the properties of this kind of situation. The best clue here is the fact that 'unusual' repetition is a *first* vocal response to any particular question. It occurs in that temporal phase when the child's attention is being drawn into the world of question and answer. By frequently not answering at all the child evades entry into this world; through 'unusual' echoes the child accords significance to what the adult has said simply by repeating it, by, in effect, saying that this is all he is willing or able to do.

REFERENCES

Atkinson, J.M. and Heritage, J. (Eds) (1984) *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press.

Baron-Cohen, S. (1989) Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology* 7. 113-27.

Brown, R. (1973) *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Casby, M.W. (1986) A pragmatic perspective of repetition in child language. *Journal of Psycholinguistic Research* 15. 127-40.

Couper-Kuhlen, E. (1989) Speech rhythm at turn transitions: its functioning in everyday conversation. Part I. *KontRi*, Arbeitspapiere Nr 5, University of Konstanz.

Couper-Kuhlen, E. (1990) Speech rhythm at turn transitions: its functioning in everyday conversation. Part II. *KontRi*, Arbeitspapiere Nr 8, University of Konstanz.

Couper-kuhlen, E. and Auer, P. (1988) On the contextualizing function of speech rhythm in conversation: Question-answer sequences. *KontRi*, Arbeitspapiere Nr 1, University of Konstanz.

Fay, W.H. (1988) Infantile autism. In D. Bishop and K. Mogford (Eds) Language *Development in Exceptional Circumstances*. Edinburgh:Churchill Livingstone.

Frith, U. (1989) Autism: Explaining the Enigma London: Blackwell)

Goffman, E. (1979) Footing. *Semiotica* 25. 1-20.

Greenfield, P.M. and savage-rumbaugh, E.S. (1993) Comparing communicative competence in child and chimp: the pragmatics of repetition. *Journal of Child Language* 20. 1-26.

Levinson, S (1983) *Pragmatics*. .Cambridge: CUP.

McTear, M.F.(1978) Repetition in child language: imitation or creation? In R.N. Campbell and P.T. Smith (Eds) *Recent Advances in the Psychology of Language*. New York: Plenum.

BEST COPY AVAILABLE

Ninio, A. and Bruner, J. (1978) The achievement and antecedents of labelling. *Journal of Child Language*, 5. 1-15.

Paccia, J.M. and Curcio, F. (1982) Language processing and forms of immediate echolalia in autistic children. *Journal of Speech and Hearing Research*, 25.42-47.

Prizant, B.M. and Duchan, J.F. (1981) The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders*,46. 241-9.

Roberts, J.M.A. (1989) Echolalia and comprehension in autistic children. *Journal of Autism and Developmental Disorders*, 19. 271-81.

Rydell, P.J. and Mirenda, P. (1991) The effects of two levels of linguistic constraint on echolalia and generative language production in children with autism. *Journal of Autism and Developmental Disorders*, 21. 131-57.

Schopler, E., Reichler, R.J., Devellis, R.F. and Daly, K. (1980) Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10. 91-103.

Schopler, E., Reichler, R.J. and Renner, B.R. (1986) *The childhood autism rating scale (CARS)*. New York: Irvington Publishers, Inc.

Sigman, M., Mundy,P., SHerman,T. and Ungerer, J. (1986) Social interactions of autistic, mentally retarded and normal children and their caregivers. *Journal of Child Psychology and Psychiatry*, 27. 647-56.

Snow, C. (1986) Conversations with children. In P. Fletcher and M. Garman (Eds) *Language Acquisition*. Cambridge: Cambridge University Press.

Steffensen, M.S. (1978) Satisfying inquisitive adults: some simple methods of answering yes/no questions. *Journal of Child Language*, 5. 221-36.

Tarplee, C. (1993) *Working on talk: the collaborative shaping of linguistic skills within child-adult interaction*. Unpublished DPhil Thesis, University of York.

Wootton, A.J. (1989) Remarks on the methodology of conversation analysis. In D. Roger and P. Bull (Eds) *Conversation: an interdisciplinary approach* . Clevedon: Multilingual Matters.

THE NATURE OF RESONANCE IN ENGLISH: AN INVESTIGATION INTO LATERAL ARTICULATIONS*

David E Newton

University of Edinburgh

1. Introduction

This paper presents an instrumental study into the nature of clear and dark sounds in English. 'Resonance' is a term which I shall be using to cover the range of quality distinctions covered by the terms 'clear' and 'dark' (and intermediate varieties).[1] The term 'resonance' has been used by a number of linguists in the past (see, for example, Abercrombie 1936, Allen 1953, and Jones 1956), as well as more recently (Kelly and Local 1986). However, its use as a phonetic label is far from universal.

2.1. The Nature of Resonance

The instrumental study detailed here will primarily look at those resonance features which are associated with the lateral consonant /l/ in

---

[*] Most of the work detailed here was carried out whilst at the University of York. The author can currently be contacted at Department of Linguistics, University of Edinburgh, Adam Ferguson Building, 40 George Square, Edinburgh EH8 9LL, UK, email den@ling.ed.ac.uk. The author wishes to thank John Kelly, Geoff Lindsey, John Local and my informants for all their advice and help. I'm still to blame though.

[1] *Clear* and *dark* are not the only terms used to refer to these particular kinds of articulatory and acoustic events. Corresponding terms in the literature include the following: 'front', 'palatalised', 'having front vowel resonance'. Similarly, terms referring to darkness include: 'back', 'velarised', 'retracted', 'having back vowel resonance', 'pharyngealized'. In the main, these terms tend to refer to the same kinds of articulatory gestures. Some of these labels may be seen as more appropriate than others, although this is not an issue which is to be confronted in this paper.

York Papers in Linguistics 17 (1996) 167-190
© David Newton

English. However, before the study and its results are described, there are three main concepts which are to be assumed in my treatment of resonance features.

Firstly, at least in their phonetic forms, dark and clear sounds are not simply opposed in a binary way. They are merely convenient labels for the opposite ends of a continuous range of distinguishable phonetic qualities. Of course, it may be the case that, when carrying out subsequent phonological analysis, one might wish to talk about a dark/clear opposition, but it is also important to recognise the range of *phonetic* variability which can be recorded.

Secondly, it often appears to be assumed in much of the literature that *clear* and *dark* are terms which only apply to the lateral consonant /l/. This has become an especially widespread assumption in that part of the literature which concentrates on the phonetics and phonology of English (see Giegerich 1992). However, work on other languages (for example, Westerman and Ward 1933), and more detailed works in general phonetics (see Jones 1956) have recognised resonance characteristics as applicable to any speech sounds.

The third point is the notion that the darkness or clearness of a token applies only to that token in a given utterance. However, upon closer examination, it can be seen that this is not the case. There have been studies suggesting that different phonetic items may have different effects on the resonance of their environment, depending on the nature of what is sometimes called their *acme function*. For instance, one study by Kelly (Kelly 1989; see also Kelly and Local 1986) examined the following two sentences, as spoken in one variety of English (from north Manchester/Salford):

(1) Ballet came to my mind.
(2) Barry came to my mind.

Electropalatography showed that the velar closure at the beginning of the word *came* was fronter following the word *ballet* than it was after the word *Barry*. Kelly proposed that, for this variety of English, /r/, in the form of an approximant, has acme function, affecting nearby parts of the utterance.

This interaction of resonance effects has also been noted by Klatt, who stated, with regard to speech synthesis, that

> 'the acoustic properties of /ɪ/ in a word like *will* cannot be predicted from diphones obtained from *with* and *hill* because the /w/ and /l/ velarise the /ɪ/ to a greater extent.'
>
> (Klatt, quoted in Kelly and Local 1986: 304)

There is also a fourth aspect of resonance features which will be discussed later on. This is the suggestion that dark tokens tend to occur finally, whilst clear tokens tend to occur initially. This is one of the more important, if problematical, aspects of the theory of resonance that is being investigated here, and will be discussed towards the end of the paper.

2.2 The Perception of Resonance

The major finding of Newton (1993) was related to how we perceive different types of resonance. That study, which was suggested by casual observations, used synthesised intervocalic laterals in English words and pseudowords produced using the YorkTalk speech synthesiser (see Ogden 1992) It was found that phonetically-trained subjects tended to perceive longer lateral tokens as having a darker resonance simply as a result of their duration, and regardless of their actual darkness or clearness. Similarly, shorter laterals were consistently judged as having relatively clear resonance, even though no differences other than duration were present.

The results obtained from this experiment raised the question of whether, for naturally-produced English laterals, darker varieties were, indeed, longer in duration. The present paper reports an instrumental study into this question, with special reference to initial and final positions.

3. Instrumental Study

This study used speech elicited from a number of informants, each being a speaker of a different variety of English.

It was found that the cue of duration in laterals seems to be of great importance in the perception of different degrees of resonance. It was hypothesised that this is because the actual duration of laterals in natural speech does indeed correlate with the resonance of the sound. Specifically, it would be expected that one would find that tokens of /l/ which are marked as relatively *dark* in a named variety are of a longer duration than those which are treated as *clear*.

3.1 Informants Used

All of the informants were first-year undergraduate students in the Department of Language and Linguistic Science at the University of York, with the exception of Speaker D, who is a member of staff there.

Four male speakers were used, each being a native speaker of a different variety of English. Their details (summarised below) were obtained through interview with each of the informants. They were also given a brief questionnaire about their linguistic background to ensure that these details were as accurate as possible:

Speaker A: 19 year-old male from Ashby-de-la-Zouch, Leicestershire, but has lived in a variety of other places. Not an RP speaker, but his idiolect is a fairly standard variety, somewhat influenced by northern English.

Speaker B: 19 year-old male from Bolton, Greater Manchester. States that he has a 'northwest Lancashire' accent, which is 'discernibly different from the more rhotic Lancashire accents (north and west of Bolton), and the Mancunian type accent which is east and south of Bolton', and is said by him to be a typical Bolton accent.

Speaker C: 19 year-old male from North Antrim, Northern Ireland. Judges that he has a North Antrim accent, but that his variety of it is not completely typical, in that his speech is 'a little more refined than where I come from'.

Speaker D: 48 year-old male originally from South-West London. Has an RP-like accent, and judges his accent to be 'RP-ish. Home Counties middle middle class'. Has lived in several other areas, but judges his accent as typical of his original background.

The use of all male speakers in this small-scale study was to make cross-subject comparisons less difficult during the instrumental study. Due to the configuration of the hardware and software, computer analysis of speech wave spectrograms is often said to be difficult for female speech, and so this was not attempted here. It should be noted, however, that in the perception experiment reported in Newton (1993) the subjects were of a rough split between female and male.

It was first hypothesised what resonance patterns speakers would have from their idiolect background, and these hypotheses were evaluated as part of the instrumental study. It was hoped to obtain the following resonance patterns for their respective articulations of /l/.

	Initial /l/	**Final /l/**
Speaker A	clear	dark
Speaker B	dark	dark
Speaker C	clear	clear
Speaker D	clear	dark

Speakers A and D have the kind of resonance patterns that are generally reported in the literature on the phonetics and phonology of (RP) English. Speaker B has what shall be called a *dark everywhere* variety, whilst Speaker C has a *clear everywhere* variety.

If it is to be assumed, following Newton (1993) and Ogden (1992), that for all speakers word-medial varieties of /l/ are of an intermediate variety with regard to their resonance, then we might expect the mean darkness (and mean duration, if the hypothesis that darker tokens of /l/ are durationally longer is true) to be classifiable into the following order (in ascending order, from clearer and shorter to darker and longer):

Speaker C → Speakers A and D → Speaker B

For the differences between Speakers A and D, it was expected that this should be in the order of

Speaker D → Speaker A*

which is possibly due to the latter's general Northern English influenced speech. These claims will be investigated below.

Some further recorded materials were also used in this study. These included some tape recordings of speakers of different varieties of English producing various utterances involving /l/ and /r/ in different environments and were recorded by Kelly and Local as part of their research work on resonance (Kelly and Local 1986). These were not used here as primary material for the instrumental study, but impressionistic observations made from them were noted for purposes of comparing results with this present study.

3.2 Utterances Elicited

The informants were asked to read out a total of 27 utterances, each of them in the form of a short phrase or sentence. The utterances were as follows:

1 say silly again
2 say sillow again
3 say solly again
4 say sollow again

5 it's the whale edition
6 the whale and the shark

7 say boy again
8 say boil again
9 say boiling again
10 say Boyling again
11 say the boy Ling again

12 say May again
13 say mail again
14 say mailing again
15 say Mayling again
16 say May Ling again

17 Mr B Likkóvsky's from Madison
18 Mr Beel Hikkóvsky's from Madison
19 Mr Beau Lukkóvsky's from Madison
20 Mr Bole Hukkóvsky's from Madison

21 Mr Beelik wants actors
22 Beel, equate the actors
23 the beelic men are actors
24 I gave Beel equated actors
25 the beeling men are actors
26 the beel equipment's amazing
27 Beel equates the actors

Utterances 1-4 are for the purpose of obtaining articulations of the same stimuli that were used in the previously mentioned perception experiment.

Utterances 5 and 6 are also examined by Halle and Mohanan (1985). These were elicited here to examine how the darkness or clearness of the articulations varies with relation to morphological boundaries.

The two similar groups of Utterances 7-11 and 12-16 were devised for the purpose of seeing how darkness varies with syntactic and morphological differences. The words *mail* and *boil* should, at least for speakers A and D, be relatively dark, as should be words *mailing* and *boiling*, since the /l/ portion is still morpheme-final. However, for the words *Mayling* and *Boyling*, one might expect a clearer articulation, since the /l/ in each case can be argued to be ambisyllabic, that is to say, belonging exclusively neither to the first syllable nor to the second, with no morpheme boundary. (For argumentation on this subject, see Local 1995.) These words should be in contrast to *May Ling* and *boy Ling*, in which it would be expected that there would be a clearer articulation. (Utterances 7 and 21 were used for purposes of comparison only, since they contain no lateral articulations.)

The remaining, somewhat unusual, utterances were all used in Sproat and Fujimura (1993). For Utterances 17-20, all the contexts were trochaic in nature (*i.e.*, a stressed syllable followed by an unstressed one), the first two being in a /i - ɪ/ environment, whilst the second two were in a /o - ə/ environment. The /l/ in Utterances 17 and

19 were made syllable-initial by the nature of the words involved, whilst those in Utterances 18 and 20 were necessarily syllable-final since they were followed by an /h/. This, as Sproat and Fujimura say,

> 'cannot be part of an initial consonant cluster in English and there is therefore no chance of resyllabification.'
>
> (*ibid*)

They also mention that, since /h/ can be considered a voiceless vowel (see Catford 1977), the choice of this sound means that there is less likelihood of interference with the lingual articulation, though they note that the laryngeal gesture for /h/ may have some side-effects.

Since the remaining utterances (21-27) were primarily concerned with drawing distinctions related to different types of morphosyntactic boundaries, these were not examined in great detail. The previous utterances (1-20) were found to provide sufficient data to be able to draw some satisfactory conclusions. However, they were examined for purposes of overview and comparison, and I shall therefore also describe them here.

Utterance 21 is similar to Utterances 10 and 15, in that the /l/ is intervocalic with no boundary, which, using the theory preferred here, is to be interpreted as ambisyllabic. Utterance 22 places the /l/ before an intonation boundary, as defined by Beckman and Pierrehumbert (1986). Utterance 23 places the /l/ before a '+' boundary, which, in Lexical Phonology (see Mohanan 1986), is a Stratum I boundary, whilst Utterance 24 places the /l/ before a phrase break within a VP.

The boundary before which the /l/ occurs in Utterance 25 is what Sproat and Fujimura call a '#' boundary (Lexical Phonology's Stratum II boundary), whilst the boundary in Utterance 26 is between the two phonological words in a compound. Finally, Utterance 27 is defined by Sproat and Fujimura as placing the /l/ before a VP phrase break.

3.3 Method

Each of the four informants was asked to read out the list of utterances in the same order. This order was chosen in a semi-random manner before the recording, so that related sentences did not appear next to each other.

Informants were given ten minutes to look through the utterances, which were written on individual cards, in order for them to be familiar with what they were going to have to read out. This was especially important, since many of the utterances are of an unusual nature, and it was important to minimise any possible pronunciation errors (though this was not completely successful; see below). The informants were each told to read the cards in a natural, but careful, style. That is to say, they understood that they were to be read as individual sentences, as this was a reasonably formal scenario, but that they should not change their accent in doing this. The recordings were later judged by members of the Department who know the informants, and these instructions were deemed to have been successful.

The recording was carried out in a sound-damped recording studio environment. The recordings were sampled into a Macintosh II computer running MacSpeech Lab II version 1.7 speech analysis software. Some of the work was later carried out by transferring the files to Signalyze (version 1.40) format, running on both a Macintosh Quadra 950 and a Macintosh LCII, though most of the analysis work was done on the former system.

Much of the analysis carried out took the form of measuring durations, and by reading wide-band spectrograms, though non instrumental techniques were also used.

4. Results

It was stated earlier that attempts to reduce misarticulations were successful, though not entirely so. Some of these did not seem to have any effect on the portion of the utterance under study. Speaker C sometimes mispronounced the word *Madison* as /ˈmeɪdɪˌsən/. In addition, Speakers A and C both pronounced *Beel, equate the actors* with less of an intonation boundary than had been intended. Again, since there was no reliance on the detail of this particular utterance, this did not cause any major problems in analysis. Of perhaps slightly more importance, Speaker C also pronounced *Beelik* as /bə'lik/, rather than /ˈbilɪk/.

Marking the start and end point for the acoustic realisation of a segment /l/ is not a straightforward task, in the sense that there are no real *start* and *end* points for the sound. Hence, two sets of measurements were made for each of the articulations. They were:

- the minimum extent where it can be said the articulation occurs,
- the maximum extent where it can be said the articulation occurs.

An example follows. On the opposite page, the display is of a wide-band spectrogram of the word *Boyling*, as edited out of the phrase *say Boyling again*, spoken by Speaker A. The two parallel sets of vertical lines show the maximum and minimum points for my measurement of the /l/ portion. These were chosen using both visual and auditory methods. This gives two different values for the duration of the /l/, depending on which criteria one wishes to use to measure it. It is therefore possible to have two discrete sets of measurements. If these both lead to the same conclusions, then there is more motivation for treating these results as accurate. In addition, these results were averaged out, to create a value for the mean length of /l/.

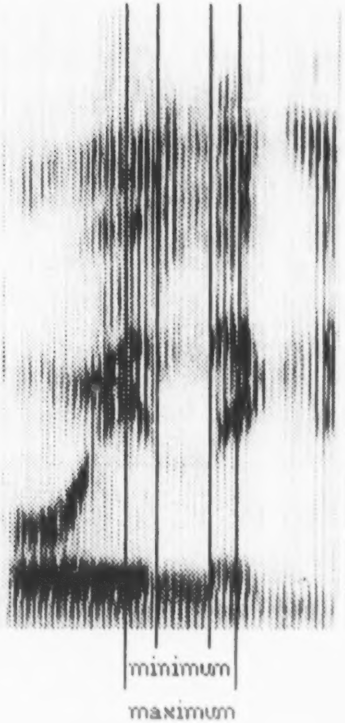
4.1 Tempo

A checking experiment was also carried out, following Kelly *et al* (1966), in order to make sure that the results were comparable across speakers. This was in relation to the tempo of the utterance. If it can be shown that the speakers' tempi are comparable (or even if they go in an opposite direction to that example given above), then we can more safely talk about the significance of any durational results that are found.

Firstly, the whole utterance was measured for each of the 27 utterances as spoken by each of the four speakers, and the total duration was noted. Secondly, a selected foot from each utterance was measured.

There were, perhaps not surprisingly, a few utterances where the tempo differences between speakers were quite large, but it was found that these differences had little impact on the overall results. The mean totals of the measurements were as follows:

	A	B	C	D
Mean length of utterance (ms)	1533	1389	1531	1440
Mean length of portion (ms)	533	485	515	540



These tempo measurements were not found to be significantly different across speakers.

It is interesting to note that, in both cases, the fastest mean tempo was from those utterances produced by Speaker B. This is the speaker who, it was hypothesised, would have longer /l/ tokens, because he had darker /l/ tokens. The fact that his speech rate was the fastest amongst the informants might suggest that, if resonance and the duration of its features was not a factor, then he would actually have *shorter* /l/ tokens than the other speakers. Hence, it is possible to say that, if the hypothesis that he had longer /l/ tokens were upheld, then this would be all the more noteworthy.

4.2 Evaluation of Resonance Patterns

The first task was to find out whether the predicted and actual resonance patterns matched. This was done partly through examination of spectrograms, but also from listening and detailed impressionistic phonetic transcription.

It was found that the resonance distributions were largely as expected. Speakers A and D had the RP-like distribution of clear initial /l/ tokens and dark final /l/ tokens. Speaker B had dark tokens in all positions (in fact, his clearest tokens were still somewhat darker than the darkest ones produced by Speakers A or D), whilst Speaker C had very clear tokens in all positions, though some of the tokens were slightly unusual, in that they did not appear to be typical of any kind of /l/ that has been discussed in this study. Some of his /l/ tokens, particularly the intervocalic ones, were very vocalic in nature, and difficult to measure. In addition, some other intervocalic tokens which he produced were tap-like in nature.

For those speakers who had a *clear everywhere* or a *dark everywhere* distribution, their final tokens of /l/ were, relatively speaking, still darker than the initial ones. Therefore, I would suggest that the RP-type classification of /l/ as 'clear initial, dark final and intermediate medial' holds at least for all the speakers under examination here, but in a relative sense.

4.3.1 Durations: Intra-Speaker

It was expected that the degree of resonance present in the /l/ part of the articulation would be classifiable in the following order, from darkest to clearest:

8 boil
9 boiling
10 Boyling
11 boy Ling

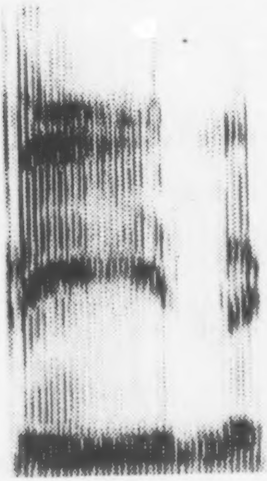
and

13 mail
14 mailing
15 Mayling
16 May Ling

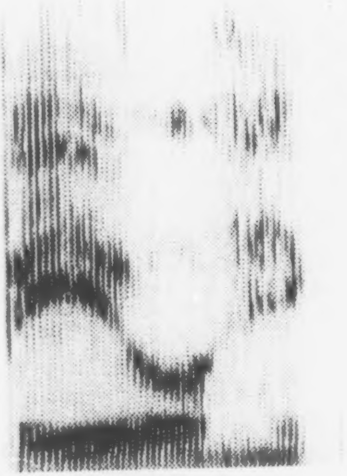
This was broadly found to be the case. For the Utterances 8-11, this pattern was found decisively for Speakers A and C, whilst, for Speaker B, the pattern was the same except that Utterances 8 and 9 were difficult to distinguish in terms of their resonance. There was an equally good result for Speaker D, with the exception that his articulations of Utterances 9 and 10 were not easily distinguishable from each other.

Similar results were found for Utterances 13-16. All speakers had the expected resonance patterns, with the exception of Speaker A's production of *Mayling*, which seemed to contain a darker /l/ than his production of *mailing*. There was a possible problem with the 'expected clear' articulations of *May Ling*. Some of these, on spectrographic study, looked as if they were in fact darker than some of the articulations of *Mayling*, even though the reverse was expected. Note that the syllable initial position of the /l/ here is in a position which encouraged primary stress location, whereas, in all the other articulations, the second syllable is an unstressed one.

This problem was avoided in Utterances 17-20, in which the expected clear articulations *B. Likkóvsky* and *Beau Lukkóvsky* are contrasted with the expected dark articulations *Beel Hikkóvsky* and *Bole Hukkóvsky*, whilst the pattern of stressed and unstressed syllables is not disrupted, as may have been the case for the articulations *boy Ling* and *May Ling*. In all cases, the expected clear articulations were found to be obviously and substantially clearer than the expected dark ones. This can be seen in the following two spectrograms (over), which are both from Speaker A.



B. Li
(from B. Likkóvsky*)*



Beel Hi
(from Beel Hikkóvsky*)*

The main visual difference between these two spectrograms is the difference in the second formant. The darker variety, on the right, has F2 falling to a much greater extent than occurs in the clearer articulation. Also, the third formant follows a similar pattern to the second in the clearer articulation, whilst, in the darker variety of lateral shown here, it moves upwards, away from the second formant. Differences in amplitude of F1 are also visible.

It was mentioned earlier that some of the informants' articulations of various /l/s were not easily recognisable. This was especially the case for Speaker C. Some of his intervocalic varieties were very vocalic in nature, making them quite difficult to segment satisfactorily. His initial varieties were also sometimes quite tap-like in nature. Speaker D produced some intervocalic articulations of /l/ which were quite fricative in nature. However, this did not cause any particular measuring difficulties.

Having ascertained that the resonance distribution was as expected, it was possible to investigate whether or not the durations of these /l/s were in a predictable distribution with the resonance.

By averaging the durations for all speakers, and including both the 'minimum length' measurement and the 'maximum length' measurement, it was found that, in the main, the results were as expected. That is to say, those articulations of /l/ which were darker in resonance also had a longer duration.

boil	70.5
boiling	54.25
Boyling	52.75
boy Ling	*62.5*
mail	79.5
mailing	50
Mayling	46.625
May Ling	*75.125*
Beel Hikkóvsky	77.5
B. Likkóvsky	60.875
Bole Hukkóvsky	85.25
Beau Lukkóvsky	73.375

In the case of the italicised articulations, the reverse durational effect has occurred. However, it was found that this unexpected effect was due to the difference in stress patterning (see above), and these results were discarded. It was then possible to directly compare the top two sets of measurements with the bottom two pairs of utterances which avoid this problem. Doing this, we find that the results are as expected, with the darker varieties appreciably longer than the clear varieties.

If the results are considered for each individual speaker, or for each of the two measuring methods, the results are not quite so consistently in favour of the hypothesis. However, no one informant's results went consistently against the hypothesis.

4.3.2 Durations: Inter-Speaker

The next piece of analysis to be carried out was to find out whether those speakers with a generally darker variety generally have longer /l/s, and whether the reverse is the case for those speakers who have an clearer variety.

The first results which were obtained were derived from averaging out *all* of the measured utterances across each speaker, regardless of in what position the lateral occurred. They were, however, not as hypothesised:

	Mean duration of /l/ (msec)
Speaker A	60.556
Speaker B	70.860
Speaker C	61.889
Speaker D	66.368

Where we would have expected Speaker C to have the shortest durations and Speaker B to have the longest, with Speakers A and D somewhere in the middle, we find that Speaker C has an intermediate value. This aside, the other speakers results are as expected, with Speaker B having an appreciably longer mean duration of /l/.

On finding this unexpected result, referral was made to notes which were made during the instrumental study. For the *B. Likkóvsky* set of four utterances, the articulations of /l/ produced by Speaker C were very difficult to segment (see above). These are the ones which, it had been noted earlier, seemed very tap-like (or at least, certainly non-lateral) when under spectrographic and impressionistic study. As a result of this, it was decided to measure these averages again, but this time leaving out these problematical four utterances. The results which were obtained this time were as follows:

	Mean duration of /l/ (msec)
Speaker A	62.429
Speaker B	70.106
Speaker C	52.210
Speaker D	64.259

It can be seen that the means for Speakers A, B and D remained almost the same, but that for Speaker C decreased by fifteen per cent. This resulted in the distribution being as originally anticipated, with the three groups of speakers (those with a generally clear pattern, those with a generally dark pattern, and those with a mixed pattern which averages out as central) each being separated by a substantial amount, around ten milliseconds in each case.

Since the laterals which were in the original perception experiment were all ambisyllabic intervocalic varieties, the means of those utterances which involved this variety of /l/ were measured for comparison. These utterances were the ones which contained the following articulations:

- silly
- sillow
- solly
- sollow
- Boyling
- Mayling

The mean durations for these /l/s were as follows:

	Mean duration of /l/ (msec)
Speaker A	52.30
Speaker B	69.83
Speaker C	46.58
Speaker D	53.00

Once again, these results were in line with what could be predicted from the results obtained in the perception experiment.

5. Summary

The results given in the above two sections support the hypothesis that darker tokens of /l/ have a greater duration than clearer tokens. This appears to be the case both for individual speakers, and also between speakers who have different resonance distribution patterns.

Some caution may be required here. I would not like to suggest that this pattern is always consistent, since the effects on resonance of morphosyntactic boundaries and their interaction with vocalic environment (and, for instance, whether one of these two factors is prime over the other) does not, as yet, appear to be sufficiently understood. In fact, as I have mentioned, some of the initial results did go against what was expected, but, to a far greater extent, the hypothesis was supported.

6.1 Discussion

One question that has been raised is whether, in general, dark tokens (of anything) are (relatively) long. Of course, since the darker varieties of /l/ which were looked at were mostly those in a final position, it is also possible that final tokens are long, regardless of whether they are dark or not. Similarly, those varieties of /l/ which were clearer were usually those which were in initial position, and there is the question of whether this is the nature of the /l/, or the nature of the position within the word, or a combination of the two. The results which were found can be schematised thus:

For /l/ only	**Initial**	**Final**
Speakers A and D	Short Clear	Long Dark
Speaker B	Long Dark	Long + Dark +
Speaker C	Short Clear	Short – Clear –

Here, a '+' sign represents that there is 'more of' the quality indicated, and a '–' represents 'less of' that quality. The actual labels themselves (*clear, dark, long, short*) represent the classifications that we might wish to give phonologically, whilst the additions of '+'s and '–'s are of a more phonetic nature.

It can be seen from the above diagram that all speakers, phonetically, do go in the same direction in terms of the durational features of their /l/s. For the order Initial→Final, all Speakers would have the order Short→Long.

This question of the possible lengthening of final items is raised by Vaissière (1983). She categorises *Final Lengthening* as a 'language-independent prosodic feature', giving examples of several languages which display this phenomenon, including French, English, German, Spanish, Italian, Russian and Swedish. However, as Vaissière admits (1983: 60), it may be too much of a generalisation to state this a universal, since there is contrary data for several languages, including Finnish, Estonian and Japanese.

If Final Lengthening could be shown to be, if not a universal, then at least a tendency, then one might wonder if there were any physiological or other reasons why this might happen. Vaissière reports several suggestions that have been hypothesised by various studies. She mentions that there may be a general relaxation of speech gestures toward the end of utterances and that this decrease in amplitude may be compensated for by increasing the duration. However, this seems to me to be a strategy that is more likely to be language-specific (or, to be more precise, dialect-specific), since we have the above examples where it does not occur. In fact, Vaissière notes that there have been studies of children, who seem not to display the tendency of final lengthening, thus suggesting that this is a learned process.

6.2 Further Study

Two areas would be relevant for investigation. Firstly, it would be interesting to find a language variety where laterals were clearer and shorter in final position. Secondly, and more generally, it would be helpful to find a language variety where non-lateral tokens were notably shorter finally than initially (perhaps regardless of resonance).

Some of these possibilities may be true for some Scottish dialects. Work carried out by the Scots Section of the Linguistic Survey of Scotland (see Hill 1960; also Hill p.c.) has suggested that some dialects of Scottish English may have very clear final tokens of some alveolars, nasals and plosives. These dialects and their resonance patterns are now being researched. If it transpires that these claims are true, it would be interesting to examine the durational properties of these sounds. If they were found to be relatively long, then this would add support to the Final Lengthening cause, whilst, if they turn out to be short, this would support the suggestion of a darkness/length correlation. In fact, preliminary non-experimental observations suggest that the latter may be the case.

I suggested above that final lengthening of /l/, if not a universal, may be a tendency. If it is assumed for the moment that this is the case, then it is then necessary to look for possible explanations. If longer tokens of /l/ usually coincide with articulations of a darker resonance, there are some reported physiological reasons why this may be so. Amerman and Daniloff (1977) studied lingual coarticulation, though they do not explicitly link dorsal gestures with increased length. However, they do suggest that the gesture of the tongue apex is the more important of the two, and that the dorsal position generally, in terms of anticipation of vowels, 'does not need to adopt so specific a position' (1977: 112). It seems possible that dorsal gestures generally take longer to activate, particularly since this would seem to involve more muscular activity, and this would be a possible explanation for the lengthening of dark tokens. That is to say, dark tokens (in this case, of /l/) have a more prominent dorsal component and dorsal components may inherently require a longer articulation period.

If this last suggestion is, indeed, a reasonable one, then it would seem to remove the need for the use of the concept of Final Lengthening, since, rather than talking about the lengthening of final items, what is here being talked about is the lengthening of the *dorsal* (or dark) items.

If further study supports this, then this would seem to tie in well with recent work carried out by Sproat and Fujimura (1993). They model all articulations of /l/ as having an apical gesture, which is consonantal in nature, and a dorsal gesture, which is vocalic in nature. One difference which they draw between clear articulations of /l/ and dark articulations is that, in clear articulations, the apical gesture occurs first, whilst, in dark articulations, the dorsal gesture occurs first. In addition, they note that

> 'the acoustically measured duration of the rime containing a preboundary /l/ correlates strongly with darkness.'
>
> (1993: 2)

They propose that the vocalic gesture has an affinity for the nucleus of the syllable, whilst the consonantal gesture has an affinity for the margin. These gestures make use of different lingual muscles. Their claim is that coarticulatory undershoot accounts, to an extent, for the correlation of darkness with duration. They also define the notion *Tip Delay*, which has a positive value in final (here, darker, tokens) and a negative value in initial tokens.

However, Sproat and Fujimura only correlate duration with resonance in the case of coda-position /l/s (1993: 18). They do not explicitly state that this is the case for all positions, nor do they suggest that this correlation is as important for perception as implied by Newton (1993). That is to say, it seems to be the case that the durational aspect of laterals may have primary status in the perception of resonance, since it has been shown that manipulation of duration affects resonance judgements when no other differences are present.

They do, however, state that their discoveries may only apply to those varieties of English which have the clear/dark distinction. They mention that there are varieties which do not display this distinction, and that there are also other languages which do not. However, of course, for the varieties used in my own instrumental study, even those which were said to be 'clear /l/ everywhere' and 'dark /l/ everywhere', were shown to have perceptible differences within these categories. As yet, I am not aware of any varieties of English having a perceptibly and consistently clear /l/ in final positions and a darker /l/ in initial positions. We do find varieties of English in which /r/, syllable-finally, is clear (for rhotic dialects, or other situations where it is pronounced), and syllable-initially is dark. However, Sproat and Fujimura do not attempt to extend their findings to any tokens other than /l/.

Nevertheless, their model could hold for other, non-lateral sounds, since the tongue gestures (as secondary articulations) for clearness and for darkness would differ in similar ways, regardless of the nature of the primary articulation.

6.3 Implications

Provided that some of the work suggested in the previous section were carried out, and that this could provide more concrete evidence for some of the suggestions presented in this paper, there would appear to be two possible implications for these findings. Firstly, there may be implications for the theory of speech production (as well as speech perception), in light of the possible clash between Sproat and Fujimura's production model and the Final Lengthening model (and in light of the perceptual findings of Newton 1993).

These findings may also have some importance in phonetic and phonological modelling, for example, in speech synthesis and speech recognition. If length is an inherent and predictable part of the structure coincident with resonance (whether this is only for laterals, or for other sounds), then it would appear to be important to ensure correct modelling of both the resonance features and these durational aspects.

REFERENCES

Abercrombie, David. (1936) Notes on the phonetics of Icelandic. Ms. University College London.

Allen, W. S. (1953) *Phonetics in Ancient India*. London: Oxford University Press.

Amerman, James D. and Raymond J. Daniloff. (1977) Aspects of lingual coarticulation. *Journal of Phonetics* 5.107-113.

Beckman, M. and J. Pierrehumbert. (1986) Intonational structure in Japanese and English. *Phonology Yearbook* 3.255-310.

Catford, J. C. (1977) *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Cutler, Anne and D. Robert Ladd (eds.) (1983) *Prosody: Models and Measurements*. Berlin: Springer Verlag.

Giegerich, Heinz J. (1992) *English Pronunciation*. Cambridge: Cambridge University Press.

Halle, Morris and K. P. Mohanan. (1985) Segmental phonology of Modern English. *Linguistic Inquiry* 16.57-116.

Hill, Trevor. (1960) Phonemic and prosodic analysis in linguistic geography. Paper presented at First International Congress of General Dialectology (Louvain/Brussels, 1960), Firthian Phonology Archive, Department of Language and Linguistic Science, University of York.

Jones, Daniel. (1956) *An Outline of English Phonetics, eighth edition.* Cambridge: W. Heffer and Sons.

Kelly, John. (1989) On the phonological relevance of some non-phonological elements. In Tamás Szende (ed.) 56-59.

Kelly, John, J. K. Anthony and Elizabeth Uldall. (1966) Tempo and transitions. Paper presented at R. I. T. Stockholm Speech Communication Seminar.

Kelly, John and John K. Local. (1986) Long-domain resonance patterns in English. *International Conference on Speech Input/Output; Techniques and Applications.* IEE Conference Publication 258.304-309.

Kelly, John and John K. Local. (1989) *Doing Phonology: Observing, Recording, Interpreting.* Manchester: Manchester University Press.

Local, John K. (1995) Syllabification and rhythm in a non-segmental phonology. In J. W. Lewis (ed.) *Studies in General and English Phonetics: Essays in honour of Professor J. D. O'Connor.* London: Routledge. 350-366.

Mohanan, K. P. (1986) *The Theory of Lexical Phonology.* Dordrecht: Kluwer.

Newton, David E. (1993) Types of resonance and their perception in English. Paper presented at the Second Manchester University Postgraduate Linguistics Conference, 13 March 1993.

Ogden, Richard A. (1992) Parametric interpretation in YorkTalk. *York Papers in Linguistics* 16.81-99.

Sproat, Richard and Osamu Fujimura (1993) Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21.291-311.

Szende, Tamás. (1989) *Proceedings of the Speech Research '89 International Conference, June 1-3, 1989, Budapest.* Budapest: Linguistics Institute of the Hungarian Academy of Sciences.

Vaissière, Jacqueline. (1983) Language-independent prosodic features. In Cutler and Ladd (eds.) 1983, 53-56.

Westerman, Diedrich and Ida C Ward. (1933, 1990 edition). *Practical Phonetics for Students of African Languages, second edition with introduction by John Kelly*. London: Kegan Paul.

PROSODIES IN FINNISH*

Richard Ogden

Department of Language and Linguistic Science
University of York

1. Introduction

Recently, it has been argued that phonetic detail ought to be accounted for by phonology: to ignore detail is to produce analyses of linguists' idealisations of data, rather than of real spoken material. Some studies of English have shown that there is phonetic detail beyond what had been expected: Zsiga (1994) has shown that post-lexical processes in English produce different kinds of [ʃ] from those produced by the application of either level 1 or level 2 rules; Manuel et al. (1992) have shown that /ð/ in English may under certain circumstances be realised by nasal portions with dental articulation and a dark secondary resonance (low F2); Hawkins & Slater (1994) show that by modelling fine details of coarticulatory behaviour it is possible to produce significantly more intelligible synthetic speech which is also more robust in difficult listening conditions. In a somewhat more theoretical vein, Docherty et al. (1995) argue that unless phonetic detail and variability is described within a phonological analysis, the analysis is seriously flawed, since it remains unaccountable to observed data. Hawkins (1995) argues that fine phonetic detail contributes to what she calls the *coherence* (naturalness) of speech. If coherence is considered important, hitherto

---

* Parts of this paper appear in Ogden (1995a). My particular thanks to Steve Harlow, John Kelly, Gerry Knowles, John Local for their help with that work. Thanks are also due to my informants, and to Tapani Salminen, who helped me decipher some of the material and produce an orthographic version of it.

York Papers in Linguistics 17 (1996) 191-239
© Richard Ogden

ignored details of speech become central properties of the linguistic system.

This paper presents a description of Finnish phonetics and a Firthian Prosodic Analysis of some of the data. Rather than starting from citation forms, the analysis is based on some of the observed phonetic detail of spontaneously produced speech.

This paper has two main sections. The first section gives a general phonetic description of my informants' speech, while the second section pays particular attention to the ways in which words in the recorded material are joined together, and presents a Firthian Prosodic Analysis of these word joins. Where the informants produce forms that are not Standard, the non-Standard forms are given in parentheses. Such forms are generally shorter than Standard forms. My impressionistic records contain as much detail as deemed necessary for the analysis presented.

The material discussed in this paper was elicited from two informants (ET and SU). Both were female, and were 17 years of age at the time of recording. They were good friends and were still at school in Kuopio, where they received instruction in Standard Finnish.[1] Since there are no substantive differences between ET and SU, utterances from both speakers are not distinguished in the text.

The material comes from two sources. The first one is a conversation between the two informants, where one describes to the other a picture so that the other informant can draw the picture seen only by the first informant as exactly as possible. The second source is a set of stories narrated by the informants based on a series of connected pictures.

My informants, who come from Kuopio, described their speech as Standard Finnish. The material elicited from them largely matches descriptions of Standard Finnish (eg. Wiik 1981, Karlsson 1982), although occasionally I obtained from my informants material which is considered typical of the Savo dialect of their home town. A linguistically trained informant from the Häme region of Finland

---

[1] Standard Finnish is a somewhat artificial language which was formalised in the 19th century. It contains elements taken from the two main dialect areas of Finnish, East and West. It is the prestige language of Finland, and the form most commonly cited by Finns to foreigners. It is also the language used in broadcasting, publishing and education.

(roughly the central south-west of Finland) identified my informants' speech as distinctively Savo on the basis of intonation. The only other striking aspects of my informants' speech in comparison to descriptions of Standard Finnish were the rhythmical structure of their words, which matches that described for the Savo dialects (Wiik & Lehiste 1968, Wiik 1975, Kettunen 1981), and their use of the glottal stop (Itkonen 1965).

2. An outline of Finnish phonetics.

My observations presented in this section are not extensive, but nonetheless provide some detail beyond commonly accepted general descriptions of Finnish phonetics[2] (e.g. Sovijärvi 1957, Wiik 1981). Notes on tempo are included, where relevant, between braces (in the manner of extIPA). Some of the standard assumptions made about Finnish pronunciation are challenged by the data in this paper. In particular, general descriptions typically do not discuss the voicing or aspiration of plosives, the precise variability in the articulation of the 'labiodental approximant' (/v/), the extent of laryngeal features such as breathiness and creaky voice, and the variability in the qualities of vowels. Standard descriptions of Finnish also concentrate on citation forms: the material on which these notes are based is not citation form, but speech produced in a relatively natural and spontaneous fashion.

2.1 Consonants with complete oral closure

Complete closure in Finnish can combine with partially or entirely voiced closure, or with voiceless closure. Complete oral closure with velic opening is only combined with voicing. The release of oral closure without nasality is generally unaspirated and the voice onset time is approximately 10-30ms (Suomi 1980, Lahti 1981). The commonest closure in normal rate speech is voiceless.

---

[2] In this paper, phonetic material is presented using an ipa font. Phonological material appears **in bold**. Orthographic material appears *in italics*.

1. {$_{\text{all}}$ ṇ măoŋ $_{\text{all}}$} kɑn:u[3]
   tämä on kannu
   this is a jug

2. jɑ nok:ɑ̆ õn tõm:ŏ̃nẽn suikŭlḛ̆
   ja nokka on tommoinen suikula
   and the spout is a kind of oval

3. no: pi:r:ĕ ʋ̥:ɑ̃:
   no, piirre vaan!
   go ahead and draw it then!

However, [k] may be aspirated, as in (4). It is not clear whether this is because it is followed by a following close front spread vowel, or whether it is because the word *kirkas* is in focal position and is pronounced relatively slowly:

4. lɑm·pŭɱ ʋɑlŏ̃:ŋ {$_{\text{len}}$ kʰirkɑs $_{\text{len}}$}
   lampun valo on kirkas
   the light from the lamp is bright

The spectrogram in Figure 7 below provides a visual of some of the phonetic characteristics of this utterance (4). Note that the first velar plosive (1) is accompanied by about 50ms of aspiration, while the second one (3) has no aspiration and the VOT is shorter, at 30ms. Note also that the apical tap (2) is voiced, not voiceless.

---

[3] Phonetic material contained between curly brackets is characterised throughout by the parameter(s) indicated subscript: {all } = allegro; {len} = lento; {p(p)} = pian(issim)o; {rall} = rallentando.

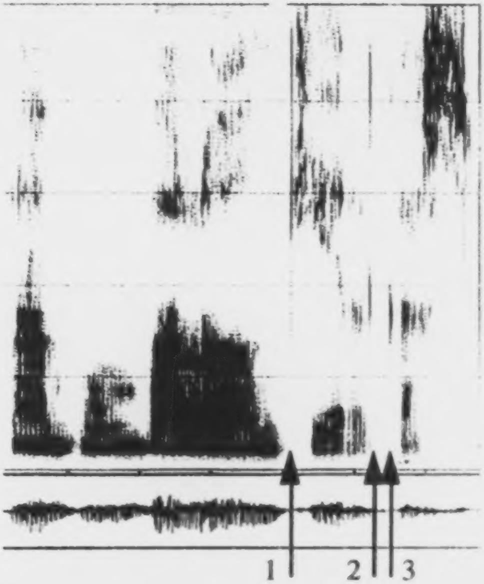


Fig. 1: [lɑmˑpŭɱ ʋɑlõːŋ kʰiɾkɑs]
'the light from the lamp is bright'

In the example in Fig. 1, the first velar plosive whose burst is at (1) is produced with aspiration and 50ms VOT, while the second one (at 3) is produced without aspiration and with VOT of 25ms, which fits in better with descriptions in the literature (Suomi 1980, Lahti 1981)

[d] occurs only in morphophonological alternation with [t]. It is articulated as a very short voiced plosive, and usually has an alveolar rather than dental place of articulation (Suomi 1980).[4] It is accompanied by a 'dark' resonance. Its closure duration is very short: usually about half the length of the voiceless plosives.

---

[4] /d/ occurs only initially in syllables which (i) contain a short vowel followed by a consonant that closes the syllable or (ii) for lexical or morphosyntactic reasons pattern in the same way (i.e. as short closed syllables) (Karlsson 1982).

5. ẽn tʲe·dă

en tiedä
I don't know

In fast speech, plosives can have a voiced closure and release when they occur in a voiced stretch of speech. Voicing with closure and release is not common word-initially. It occurs most frequently in words formed from pronouns, as in *tommoisella* in example 6, and after periods of voicing and lateral airflow:

6. $_{\{all}$ tseõ ʋɑ̃: domːozelə $_{acc}$ ʔyɦel·ä ʋiːʋăl·ă tehtÿ $_{all\}}$

se on vaan tommoisella yhdellä viivalla tehty
it's made with one sort of line

7. naytːa: korʋaldə

näyttää korvalta
looks like an ear

8. naytː$_{\{all}$ aːɣö ne ðaːldah $_{all\}}$

näyttääkö ne täältä?
do they look like this?

9. boh·jăn

pohjan
bottom (gen.)[5]

---

[5] Said as a repetition of the previous speaker; the previous utterances are recorded in example (43).

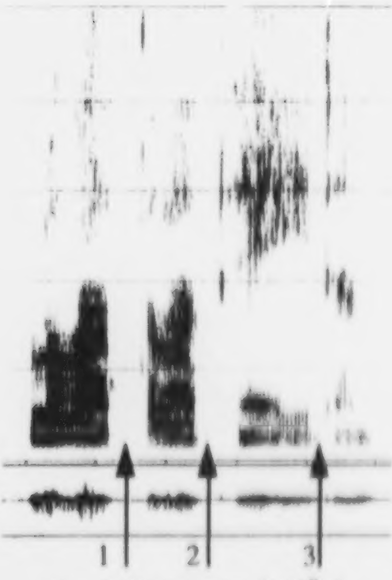


Fig. 2: [luvɑt ovɑt heidẵn]
'the licences belong to them'

Note the three different closure durations for the plosives. (1) was measured at 90ms, (2) 60ms, and (3) at 40ms. In this instance, the amount of voicing for [d] is very small, and the duration probably gives the strongest cue to the status of the plosive.

When short and in the initial portion of an unstressed syllable, plosives can sometimes be articulated with a stricture of less close than complete closure, giving [p̞ ɟ k̞] or even friction and voicing. There are insufficient instances of this in my data for it to be possible to work out whether there are any systematicities in the way this is used. However, it seems true to say that the weaker closure occurs before unstressed syllables, and only when the stretch as a whole is voiced. Closure portions are always followed by audible release within the word (where there is only one plosive-plosive cluster: [tk]). However, between words, the plosive [t] has a variety of release types. It may be released medially:

10. {$_p$ namˠ·ɑ̩t $_p$} kɑ:p:ĕjɑ̆
nämä ovat kaappeja
these are cupboards

When a lateral follows, it may be released laterally:

11. {$_p$ namḁoʋɑt^l $_p$} lɑmp:ŭjɑ̆
nämä ovat lamppuja
these are lamps

When a bilabial plosive follows, there may be no audible release:

12. hɑtut˺ pɑn:ɑ:m pa:ɦ̤ɑ̤̃n̤
hatut pannaan päähän
hats are put on the head

It may be that in the case of apical followed by bilabial closure, the bilabial closing gesture masks the release of the apical closure. In other words, the bilabial closure is timed so that it happens before the apical release.

Unreleased closure is a common way for a speaker to keep hold of a turn in a conversation. When this closure is released, the next stretch of speech sounds like it begins with a plosive (e.g. (7) above, which begins with a portion transcribed [ts-] and is preceded by [-ʔ] and a pause).

2.2 Velic opening and oral closure: [m ɱ n ŋ]

Nasality co-occurs with complete oral closure made at various places in the oral tract: bilabial, labio-dental, dental, and velar. Nasality and voicing always co-occur in Finnish. Finally in the syllable, nasal consonants are articulated homorganic with any subsequent plosive; otherwise they are articulated as apico-dentals. (See Section 3.1 **n**.) [n]

is produced with the tongue tip just back of dental and forward of the alveolar ridge.

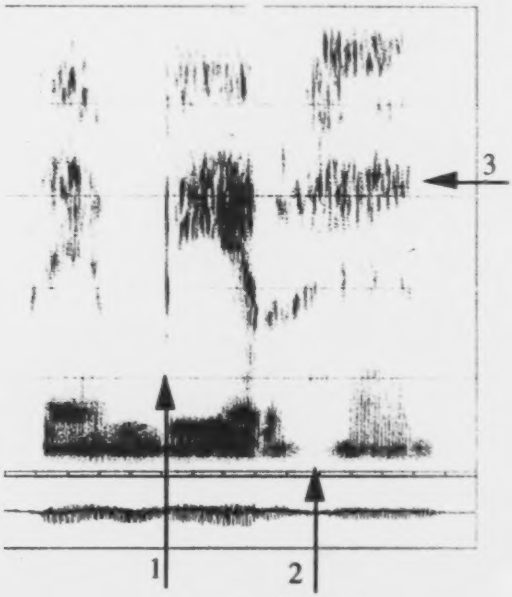


Fig. 3: [tẽ̃iŋ·ˈirɦẹ̃ː:n]
'I made a mistake'

Note how the nasal portion ends with a very obvious plosive-type release (1); the low amplitude of voicing for the initial part of the last syllable (2), and the breathiness throughout this final syllable (3).

13. {$_{p}$ e m:a ti·a mil·tă ne n:a̠yt:ạ: $_{pp}$}
en mină (mă) tiedă (tiă) miltă ne năyttää
I don't know what they look like

14. mĩŋkălɑinẽ se ʔɑlɑ·pa̠: oli·h

minkälainen se alapää oli?
what was the bottom bit like?

15. noh: tam oŋ kɑn:ŭŋ kɑu̠lɑ

no, tämä on kannun kaula
well, this is the neck of the jug

In portions with nasal and labiodental articulations, there is a great deal of variability, from apical contact with nasality to labiodental contact with nasality. In the latter case, it may be that this labiodental contact is completely coextensive with nasality, and that length together with labiodentality are the only exponents of the syllable-initial C. Release is marked with a superscript ! in (20).

16. seinăn ʋierĕs:ă

seinän vieressä
next to the wall

17. tẽĩɱ·$^{!}$irɦẽ̤:n

tein virheen
I made a mistake

2.3 Lateral airflow

Laterals are articulated dentally in Finnish. When a nasal precedes a lateral, nasality may extend into the lateral portion, and laterality and nasality may be produced simultaneously. Finnish laterals are on the whole darker than their English counterparts, but are never as heavily velarised as finally in English syllables.

2.4 Tapped and trilled articulations

Taps and trills seem to be in free variation in my informants' speech; but taps (but not trills) are in free variation with the voiced plosive [d]. Another informant (from Häme) has trills and taps where my informants have [d]. In citation forms and careful speech, the trill [r] has 2-3 vibrations of the tongue when short, and 5-6 when long. In fast speech, the tap [ɾ] counts as the exponent of 'short' and the trill has 2-3 vibrations of the tongue, and counts as the exponent of the category 'long'. Both taps and trills are pronounced voiced in clusters with voiceless plosives: [kerto:], not [ker̥to:] *kertoo*, 'tell', 3ps. present tense. Initially however they may sometimes combine with a short period of voicelessness.

18. ma onĩ pi:rtănyh
minä (mä) olen (oon) piirtänyt (piirtaäny)
I have drawn (it)

19. lafiel:ă reɰnɑ:
lähellä reunaa
near the edge

20. tom:ŏnẽŋ koɾʋă
tommoinen korva
a sort of ear

21. ʋihˑre:lˑa ʔẽ̞nsin te:t ʋar:ɐ̠t
vihreällä ensin teet varrat
you do the stalks first in green

Sometimes lateral and trill articulations are found with initial voiceless portions utterance-initially:

22. l̥ɑsˑkĕ v̩:i:tẽ:n
laske viiteen
count to five

23. r̥ɑkẽ̆ˑnsĩ^{n} tɑlon
rakensin talon
I built a house

2.5 Open approximation

Two approximants occur in Finnish: palatal and labiodental. The labiodental approximant is often accompanied by a somewhat ballistic lower lip gesture, producing something like a labiodental flap. Sometimes in the initial portion of a stressed syllable, the stricture for the labiodental approximant is that of rather close approximation, producing weak friction; it is not uncommon word-initially to hear a voiced labiodental plosive (see Fig. 3). The palatal approximant does not exhibit this wide range of variability in its degree of stricture.

Approximants only occur syllable-initially. (Flifilet 1971; Suomi 1985a and the references therein consider whether this distributional pattern is evidence for treating the final component of diphthongs, which may be [i] or [u], and initial approximants as allophones of the same phoneme.)

24. tụn˙$_{\{\text{rall}}$ tĕ̃maton lajiç$_{\text{rall}\}}$[6]
    tuntematon lajike
    an unknown species

25. no te: ʋ:ɑᵊk:ɑ ruiskuk:iɐ
    no, tee vaikka ruiskukkia
    well, why don't you do cornflowers

Sometimes in back harmonic words, the palatal approximant is very back, and is transcribed as an advanced velar glide. There are not enough instances of it in my data to be able to say anything very conclusive about it.

26. hɑp:oɰ̟ ɐ
    happoja
    acid, part. pl

---

[6] Note here that the utterance ends voiceless, as is common for utterance-finals. Note also that it is a dorsal articulation, and that it is front. It would be inappropriate to regard this as some form of deletion, since all the phonetic properties demonstrated at the end of this word can be shown to be systematic. See Section 3.6 h.

2.6 Friction with and without voicing

The fricative [s] can be produced in Finnish with the tongue tip down. This produces a rather flatter, duller sound than in, say, English. The groove is also wider than in English, enhancing this impression of dullness (cf. Sovijärvi 1957).

Another variant of [s] is also found. In this articulation, the groove made by the tongue is considerably narrower than in English, and the tongue tip is up. The groove made by the tongue forms a narrow V-shape from the blade to the tip. The result is that this [s] sounds whistly to English speakers. The data I have suggest (but not conclusively) that the whistly [s] sound occurs before front, spread non-open vowels. When these two articulations are combined with secondary articulations affecting mostly the dorsum and harmonising with the resonances of the neighbouring vowels, a gradual spectrum of qualities is produced rather than the simple two-way split suggested here. Nevertheless, the 'whistly' articulations do stand out in the recordings.

The records below show examples. The 'flat [s]' is transcribed [ṣ] and the 'whistly [s]' as [ş]:

27. ɑṣuin ṣiːnă tɑloṣːŏ
asuin siinä talossa
I lived in that house

28. lafide ɑṣe·mɑlːĕ
lähde asemalle!
go to the station

29. kɑto·ṣĩ mːet·şãːn
katosin metsään
I disappeared into the forest

The different types of [s] sound are not marked elsewhere in this paper. Between voiced sounds and within words, weak voicing may cooccur with apical friction which is of short duration:

30. nouzĕ s:ăŋ:y̆stḁ
nouse sängystä!
get out of bed

There are in the data some instances where a word begins with initial voicing and friction. These words are commonly pronouns, as in the words *täältä* (demonstrative pronoun, ablative sg.) and *tuonne* (demonstrative pronoun + illative sg.) in the examples below; strictures of relatively open approximation in fast speech are sometimes also found instead of strictures of complete closure. In these cases, the friction is rather weak.

31. nayt:{all a:ɣð̆ ne ða:ldah all}
näyttääkö ne täältä?
do they look like this?

32. {all jos si̥ta kɑ̈ðoˑt s all} ʔylha:lta pain
jos sitä katsottaisiin (katottas) ylhäältä päin
if you looked at it from above

33. jə kɑɦʊð̆ tule: {all ðone ʔoijç̬ˑle all} pwolə̥l:e
ja kahva tulee tuonne (tonne) oikealle puolelle
and the collar comes up to the right-hand side

2.7 Voicelessness, breathy voice: [̥ h ɦ]

Phonetically, it is perhaps best to see Finnish [h] as a voiceless version of an adjacent vowel. This is also Sweet's description of Finnish [h] (Sweet 1908, in Henderson (ed.) 1971: 174). 'There is also a "strong" aspirate which occurs in Finnish and other languages, the formation of which the full vowel position is assumed from the beginning of the aspiration, which is therefore a voiceless vowel.' On the other hand, the degree of aspiration at the syllable margins is greater than in the voiceless vocalic syllabics noted below.

[ɦ] can be treated in a similar way, as a breathy voice version of an adjacent vowel. [ɦ] occurs between two voiced sounds, and [h] elsewhere. Both [h] and [ɦ] are found syllable-initially and finally.

In my informants' speech, [ɦ] as a distinct portion of breathy voicing focused at the syllable margin is frequently not observed, but breathiness throughout the syllable is. This is especially interesting in view of some of the metathesis which is supposed to be fossilised in Finnish (cf. Rapola 1966: 256ff). In the Standard language, there are pairs of words such as *valhe*, 'a lie' and *valehtella* 'to tell a lie'.[7] When my informants were asked to give the word for 'a lie' they consistently produced [ʋɑl·e̤ː], with breathiness throughout the whole of the second syllable (or if anything concentrated on the latter portion of it); but certainly not initially in the syllable as the (generally phonemic) orthography implies.Note that the lateral portion of this word is pronounced half-long, where half-long duration serves as the regular phonetic exponent of the first element of a CC-cluster (cf. Ogden 1995b).

Fig. 4 presents a spectrogram a token of the word *hiihdin*, 'I skied', where the whole of the first syllable is pronounced breathy .

---

[7] Similarly, there is the word *paras*, 'best', which has the stem *parhaa*-; /h/ may not occur finally, since only apical sounds occur in this position. This instance can be seen therefore as an example of metathesis of friction.

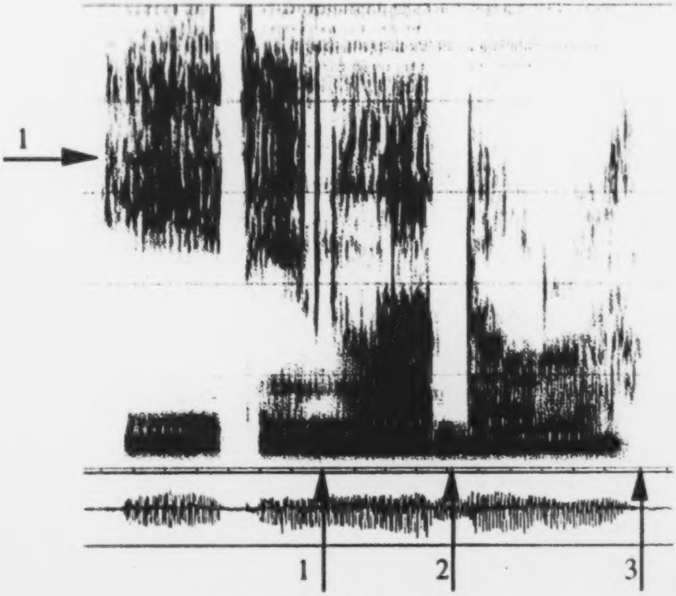


Fig. 4: [hi̤:ɦdĩn lɑdŭl:ɐ̥̆]

'I skied on the track'

Note the breathiness evident throughout the first syllable (1); the very short voiced closure for the [d] sounds (2), and the final voicelessness (3).

At the end of a syllable, the tongue gesture for the vocalic part of the syllable may be somewhat raised and accompanied by voicelessness, producing weak friction, as in [lɑχti], 'Lahti', a place name.

34. ʋi̯:ʋǫistɐ̥̆ tȩhtÿ̆jḁ̈̆
viivoista tehtyjä
made of lines

35. jɑtkɑt vaɦã
jatkat vähän
you go on a bit

Voicelessness is frequently used to mark utterance finality. Stages into complete voicelessness from voicing are typically: voicing, creak, voicelessness. Voicelessness may frequently be accompanied by quietness. Sometimes the voicelessness is rather 'strong' (recall Sweet's observations), and is then transcribed as [h], with the meaning that a more forceful articulation is used that that implied by the symbolisation using a voiceless vowel.

36. ihmẽ· l̄:isa:ke t:one ʔoikḛ:le p·wo̰:l:ḛh
ihmeen lisäke tuonne (tonne) oikealle puolelle
a strange appendage on to the right hand side

37. {$_{pp}$ e m:a ti·a mil·tă ne n:a̰yt:ḁ: $_{pp}$}
en minä (mä) tiedä (tiä) miltä ne näyttää
I don't know what they look like

38. kis:ɑ̥ ʔistui mɑto̰l:e̥
kissa istui matolla
the cat was sitting on the carpet

See also below, 'Voiceless vowels'.

2.8 Glottal stop and creaky voice: [ʔ _]

The glottal stop and creaky voice are frequently used in the speech of my Savo informants to mark the beginning of words which have a vowel initially. Lehiste (1965) presents some similar data comparing vowel-vowel sequences with and without intervening syllable boundaries; those with syllable boundaries may use creaky voice as in Fig. 5.

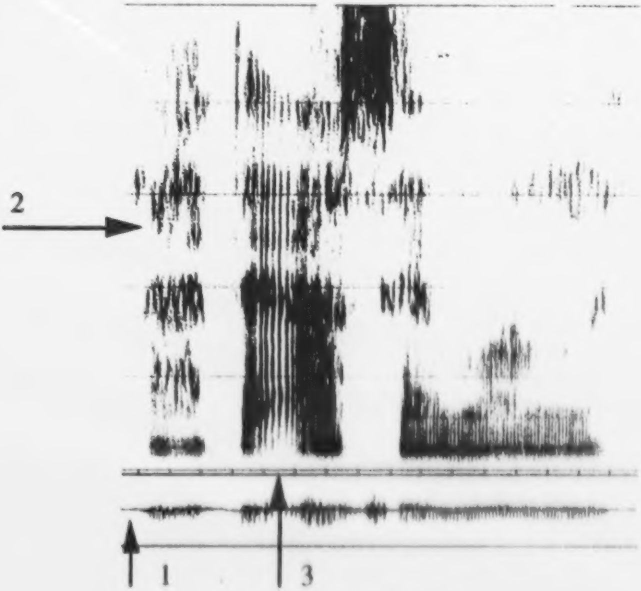


Fig. 5: [l̥a̤ɦde_ɑsẽmɑlːç̆]
'go to the station'

Note the initial voicelessness (1), breathiness throughout the first syllable (2), and the very striking creaky voice between the second and third syllables (3). Much of the transition from one vowel sound to the next coincides with the period of creaky voice.

39. õŋksĕ pyøreː seɑ ʔɑlhɑːlːɑ ʔoleˑvɑh
onko se pyöreä, se alhaalla oleva?
is it round, the one underneath?

40. ɑikɑ̆ ʔisoˑ
aika iso
quite big

41. ʔẽn:sĭn ʔotɑˑ m:usˑtŏ tus:ĭ
ensin ota musta tussi
first take the black pen

42. {f all mĩŋkălainẽ se ʔɑlɑˑ all f}pa̤: oliˑh
minkälainen se alapää oli?
what was the bottom part like?

Another function of glottal stops in conversation seems to be as a device for keeping hold of the turn in the conversation. While one speaker has an unreleased closure, the other speaker does not interrupt:

43. ma pi:rsĩn siřaˑʔ˺... tam pʸøˑrÿlă osɑˑn jɑʔ˺... jɑ pohˑjăn
minä (mä) piirsin sitä... tämän pyöryläosan ja... ja pohjan
I drew it... this round bit and... and the bottom

More detailed descriptions of creaky voice are given in Section 3.3 under the exponents of **ʔ**.

2.9 Resonance features

With the possible exception of [d], consonants in Finnish match their resonance with that of the vowel of the syllable in which they appear. However, there are not such extremes of consonantal articulation that consonants with palatal place of articulation or heavily velarised consonants are produced.[8] These seem not to form part of the Finnish repertoire.

Consonants in words with back harmony are consequently darker than in words with front harmony. One way of delimiting words is a change in the resonance of the consonants at the words' edges:

---

[8] A distantly related language, Nenets, lost 'vowel' harmony early in its development and now has palatalised and velarised consonants. Finnish secondary articulations are not as extreme as these.

44. kaytĩmʲ pʷˠufie·lintă
käytin puhelinta
I used the telephone

Note that in this example, the words are kept together by the shared bilabial place of articulation but are kept separate by the different resonances. The resonance of consonants is not marked in my transcriptions unless it is different from what is expected.

Lip-rounding, which is predictable, is similarly not transcribed for consonants, although it must be noted that the lips hold the same gesture over the whole syllable, or in the case of diphthongs over the syllable-initial or syllable-final piece.

As far as [d] is concerned, it could be that it is the low-frequency voicing during the closure which gives the auditory impression of darkness. It should be added that some writers (eg. Karlsson 1971) believe that this voiced alveolar plosive is an import from Swedish and that it came about when the modern language was standardised in the capital Helsinki in the last century—Helsinki was at that time predominantly a Swedish-speaking city. Kettunen's map 65 (Kettunen 1981) shows that [d] only occurs natively in one or two areas on the West coast, which, significantly, are also areas where Swedish has a strong foothold. My informants were able (consciously) to produce dialect forms which used other articulations than the one described here such as a voiced bilabial approximant or a voiced tap. TS, my informant from Häme, regularly uses a voiced apical tap or trill in all contexts where [d] appears in the data presented here.

2.10 Vowels

The symbols used in my records for the vowels are: [a ɑ o ø u y i e ɐ]. This follows the usual IPA practice for Finnish vowels, although the orthography is more common: <ă a o ŏ u y i e>. The symbol [ɐ] (sometimes also transcribed in my records as [ə̜] for a slightly closer vowel) is used to represent an open, central quality which is frequently found in unstressed syllables, particularly very short ones[9]. It is normally accompanied by a diacritic for advancing or retracting.

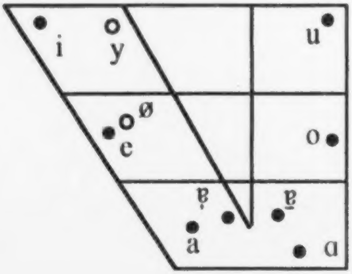


Fig. 6: Vowel quaderilateral showing the approximate qualities of Finnish vowels.

The symbols used in the transcriptions presented in this paper are used as follows: [a] is not as open and front as CV4, nor is [ɑ] as back; its quality is rather more central though very open. The mid vowels [e ø o] are all more mid in quality than their IPA symbolisation implies, though they are hardly less peripheral. [u] is very back and round, almost cardinal. [i] is front and spread. [y] on the other hand is not so front and is less rounded than, say, French [y]. It bears some resemblance to the short German sound [ʏ] as in *wünschen*. Diacritics accompanying vowel symbols modify the values described here, and not cardinal vowel values.

No significant differences in quality have been observed for Finnish vowels depending on their duration (cf. Sovijärvi 1938, Wiik 1965, Engstrand & Krull 1984).

---

[9] cf Harms (1964: 62), who uses the symbol [ʌ] for this sound in back harmonic words. He claims it appears only when preceded by a syllable boundary or following a consonant cluster, and only in or beyond the third syllable. My notes do not quite accord with this last observation, and I have observed both fronter and backer varieties.

2.11 Diphthongs

The so-called rising diphthongs of Finnish all end in a close vowel. They are: [ai ɑi oi øi ui yi ei], [ay ɑu ou øy eu] (and, marginally, [ey iu iy]). The diphthongs which end spread do not normally end as close as the symbol [i] implies: they usually fall somewhat short of this, to approximately [e] or [ẹ]. The diphthongs that end spread but which are not in the first syllable of the word are usually 'derived', ie. they are not part of the stem of the word, but arise from the addition of [i], which marks past tense and plural in Finnish.

The so-called opening diphthongs are: [uo yø ie]. These vary in their articulation depending on the speaker's dialect (Kettunen 1981). My own informants pronounced these sounds as scarcely diphthongal. They tended to start with a short close portion opening to a mid portion which nevertheless was quieter than the initial part of the diphthong, e.g. [k̟ʷo:rutet:ɐ̝] 'icing', part. sg., [tʏø:tøn], 'unemployed', nom. sg. In Standard Finnish these vowels have longer initial portions with a mid off-glide. These diphthongs are usually treated as the phonetic exponents of long mid vowels, since in the first syllable (the only place where they occur), [e: o: ø:] — ie. pure, long vowels — are only found in loan words. In native words, therefore, the long vowels are in complementary distribution with the opening diphthongs.

2.12 Velic opening and vocalic articulations

The timing of the lowering of the velum is generally such that it lowers before a complete oral stricture is made, producing vowels which are nasalised before nasal consonants. Word-finally, there is frequently no complete oral stricture, but there is audible nasality throughout the final syllable. Lehiste (1965) shows that the nasalisation of a vowel may serve as a boundary marker in Finnish. The pair *maan isä* and *maa nisäkäs* are distinguished partly by the fact that the first vowel of *maan i-* is nasalised, while in *maa ni-* it is not.

2.13 Variability of vowel quality

Vowel qualities produced by my informants are somewhat variable; this variability can be summarised somewhat, though some of the observations in this section remain rather tentative.

- Very short vowels tend to be centralised.
- Vowels after the palatal glide are frequently fronter in quality than elsewhere; but it is hard to tell whether there is anything substantial to be said here, since these vowels also tend to be very short in my data, occurring as part of the partitive plural suffix.
- Vowels after apical consonants tend to sound slightly fronter in quality than after labial or dorsal consonants.

Some examples from my data will give an impression of the kinds of variability in vowel quality which can be observed.

Compare the formant values for the centre points of the three open vocalic portions in the word [ɑm:ɑ̆t:ĕjɑ̣̆]. The first one has the formant values 855-1520-3335 Hz, and the second one 765-1570-2965 Hz. These are roughly comparable; taking into account the fact that the second one is short and occurs between two consonants, one might expect a lower F1 value; the F3-F2 difference might be explained by the proximity of bilabial closure, which tends to lower all the formant values. The final open vowel however has the formant values 815-1875-2945 Hz, which is quite a lot fronter (i.e. with a higher F2) than the other two open vowels. Bearing in mind the fact that this vowel is also very short, and also next to a palatal approximant (which would have slower formant transitions), this high F2 value might be explained by coarticulation. However, one of my informants produced the word *housujaan* 'his trousers', part. pl. as [housuja̤:n]; this makes it more likely that there may be some kind of local harmony between the palatal approximant and the subsequent vowel.

A kind of harmony may be observable within feet. The observations made here are by no means conclusive, though they are suggestive. In the phrase *pidän ammatistani* 'I like my job', it was observed that the third open vowel in the word [ɑm:ɑ̆tistɑ̣ni] was fronter than the other two open vowels (with formant values of 695-1885-3015, thus roughly comparable with the third open vowel in *ammatteja*). Three possible explanations seem likely: (1) the vowel is in a foot with two syllables with front resonance: perhaps there is vowel-to-vowel coarticulation; (2) the vowel is surrounded by apical consonantal articulations, which tend to raise F2 and so give the impression of fronter vowels; (3) the functional load on the vowel so late in the word is minimal, and no other vowel could occur in that place in structure and make a difference in meaning, therefore one might expect that this vowel would have the potential to be more variable in quality; example 57 is a similar example of this. It may also be the case that all three explanations have some validity.

2.14 Voiceless vowels

Vowels between voiceless consonants are sometimes voiceless. This seems typical of fast stretches of speech, turn ends, or stretches where as the result of metrical structure the vocalic portion would be very short even if voiced.

45. mitaˑ kuk:ɑ: ne m:uistŭ̥t:ɑ̥:

mitä kukkaa ne muistuttaa
what flower do they remind you of?

46. {all jos si̥ta kʌðoˑt s all} ?ylha:lta pain

jos sitä katsottaisiin (katottas) ylhäältä päin
if you looked at it from above

47. lɑmˑpŭ̥t oʋɑt kirk:ɑitɑ̥

lamput ovat kirkkaita
the lamps are bright

Just as certain consonants are voiced in stretches which are overall voiced, so it appears that short vowels in stretches which are overall voiceless can be voiceless.

2.15 Quantity and Duration

There are many different quantities for both consonants and vowels in Finnish. At the phonological level, it is usually said that there are two contrastive degrees of length. At the phonetic level however, it is not true to say that there are only two degrees of duration. In my records five degrees of duration are marked: [^{v} v̆ v v· v:]. Note that it is more accurate to see duration as gradient rather than as categorial, so that no matter how refined the transcription, the records remain impressionistic rather than conclusive.

Half-long vowels are found after short open syllables, giving the shape [cvcv·] (cf. in particular Wiik & Lehiste 1968, Wiik 1975, who show that the precise duration is a dialectal matter: some dialects have the shape [cvcv]). Half-long vowels in my informants' speech frequently occur also in closed syllables, provided the syllable-final consonant is a sonorant (typically [n]), giving the general shape [cvcv·n]. This pattern is not found when the final consonant is a voiceless plosive (usually [t]). [cvcvt]. Palomaa (1946) found that vowels before voiceless consonants are shorter than before voiced ones.

Half-long consonants appear as the exponent of the first element of CC-clusters, giving the general shape [cvc·cv].

Very short vowels are found after heavy first syllables, giving the phonetic shapes [cvvcv̆] and [cvccv̆]. A short vowel after such a stretch may also be very short: [ɑm:ɑ̆t:ĕjɑ̆ kɑ:pĭstɑ̆] *ammatteja*, 'profession' part. pl., *kaapista* 'cupboard', elat. pl.

Factors which may be significant in determining consonant duration are: place in the foot; the weight of preceding syllable; and the phonological length. In one token of the utterance *tapaa nainen ulkona* 'meet the woman outside', the four nasal portions had the following durations respectively: 85ms, 35ms, 70ms, 60ms.[10] The first one counts as the phonetic exponent of a 'long' nasal, while the others are 'short'; however, it can be seen that there is a wide range of variability in the measured durations. Clearly, there can be no simple phonetic interpretation of the categories 'long' and 'short'; and any interpretaion would have to make reference to position in the word, syllable, and foot. See Local & Ogden (1994) for a desription of a computationally implemented method for generating consonant durations for English in a declarative metrical framework.

---

[10] cf. Flifilet (1971) who in discussing Finnish rhythm notes that consonants after long vowels are very short.

Occasionally, my informants demonstrate a feature judged typical of their dialect: after a short open syllable, and before a long vowel, phonologically short consonants can be durationally long. This type of lengthening depends purely on the metrical structure and plays no part in morphosyntactic processes, unlike the well-known 'consonant gradation'. This is not a feature of Standard Finnish, and is not reflected in the orthography.

48. men:e: ʋaɦa ʔɑlɑˑs pa̰ĩ n:ḭ̃:ŋkṵs:a̰h
menee (mennee) vähän alas päin niin kuin sinä (sä)...
goes down a bit like you...

49. ei mit:ã:n
ei mitään
nothing

3. Inter-word Junctions in Finnish.

This Section presents a Firthian Prosodic Analysis of inter-word junctions in Finnish. Some of the phonetic facts described in Section 2 are taken account of by the analysis presented here, and more data is presented to back up the analysis.

In Fithian Prosodic Analysis, syntagmatic relations can be considered primary: one starts by considering how linguistic items are put together. This avoids the need for assimilation rules (Sprigg 1957), and may also avoid the need for deletion rules. The fundamental nature of syntagmatic relations is expressed by Whitley (ms), below:

> 'You can't tell from your isolate form what the junctions will be. You have to start from the junctions—you *can't* work from the isolates and say x *becomes* y in certain circumstances.'

Thus, for Whitley, citation forms ('isolates') do not provide the starting point of the analysis; instead, she prefers to begin with items in connection with one another. This is how the analysis of the Finnish material in this section is conducted. The resulting statement is very different from one which starts out with citation forms which have to be altered to fit in with rules of word juncture. I will also show how at least some of the observations made in the preceding section can be taken into account.

In the analysis presented in this Section, I shall assume a structure $\omega\text{-}^{\pi}\text{-}\omega$, where ω stands for 'word', and π for a system of word junctions. I shall then consider whether the terms of this prosodic system can usefully be reused in the prosodic system of syllable joins within words.

In all, there are six terms of the prosodic system of inter-word junction in Finnish: **n g h ζ ʔ τ**. As long as the stated structural constraints are not violated, up to two prosodies of word junction may operate at one place in structure; but every ω—ω structure must contain at least one prosodic term. The term is largely (but not entirely) determined by 'phonematic' structure, although lexical and morphological structure also play a part. I shall consider each kind of junction in turn, considering firstly its distribution (i.e. its phonological status), and secondly its phonetic exponents. The term **N** is used as a word-final phonematic unit whose exponents include nasality; it is a term more delicate than **C** (which merely stands for any term of the relevant **C**-system) and as delicate as **P**, which stands for a subterm of the **C**-system and whose phonetic exponents at normal tempo include complete oral closure.

The data in this Section have a different relevance from the data in the preceding Section, and are consequently presented differently. In this Section, the focus is more on the relations between the phonetics, the phonology, and other levels of linguistic statement such as the grammar. Therefore, impressionistic records annotated with the junction prosodies in bold superscript are given, along with the generalised partial phonological structure of which the phonetics is an exponent, a brief account of the morphological structure of the items, and an English gloss.

3.1 n

Distribution of **n**

n is found at the junction of two words where one word ends in -N and the subsequent word begins with a C- whose exponents include maintainable oral stricture (Catford 1988: 63) which involves the actual physical contact of an active and passive articulator; ie. the exponents of C- include [p t k m n s l r ʋ], but preclude [j h].

Exponents of **n**

n pieces are characterised by the same place of articulation across the syllable ending and the syllable beginning. The presence of nasality determines the presence of voicing, but nasality may terminate before voicing. In the case of the exponents of the structure -N $^{\mathbf{n}}$ **P**-, voicing may extend into the closure portion which is one exponent of **P**-. Nasality may occasionally extend into the syllable beginning and combine with labiodentality or laterality.

Nasality is perhaps best regarded as the exponent of -N, but the temporal extent of nasality may best be regarded the exponent of **n**.

Note that what is accounted for by **n** is accounted for in other analyses by rules of assimilation (eg. Karlsson 1982: 144). These rules assume that the base form of the word ends in /n/: when a word with final /n/ precedes a word with, eg., initial /p/, then the nasal assimilates. Such assimilation rules are only necessary because the starting point of the analysis is citation form words; these forms are dealt with under **τ** below. Furthermore, these analyses do not account for the range of variability in the exponents of pieces of the structure -N **v**-, where the exponents of **v** are labiodentality and approximation (cf. Section 2.5).

Examples:

50. mu:tɑmɑŋ $^{\mathbf{n}}$ kort:člĩm $^{\mathbf{n}}$ pa:ɦanh $^{\tau h}$
C—N $^{\mathbf{n}}$ P—N $^{\mathbf{n}}$ P—N $^{\tau h}$
{several+gen block+gen top+ill}
down a few blocks

51. nɑinẽ·ŋ̆ ^{n} kɑufiistu:
C—N ^{n} P—V
{woman+nom, is terrified+3ps}
the woman is terrified

52. mentɑʋɑ $^{\zeta}$ tɑkɑisĭŋ ^{n} goti:n $^{\tau}$
C—V $^{\zeta}$ C—N ^{n} P—N $^{\tau}$
{go+pass+pres. part back home+ill}
has to go back home

53. ǫʋen ^{n} ɱɑli:n $^{\tau}$
V—N ^{n} C—N $^{\tau}$
{door+gen through+ill}
through the door

54. ɑ̣n $^{\tau ?}$ ǫstɑ: $^{\zeta}$ kiŋ ^{n} gel:ðn $^{\tau}$
C—V $^{\tau ?}$ V—V $^{\zeta}$ C—N ^{n} P—N $^{\tau}$
{3ps+nom buy+3ps clitic clock+gen}
and he buys the clock

3.2 τ

Distribution of τ

τ occurs in several structures: (i) Wherever the first part of a junction is any term of the final -C system except -N. (ii) When any -C term (including N) is utterance-final. (iii) In the structure -N $^{\tau}$ C-, where the exponents of C- include a non-maintainable stricture, or no stricture (ie. [j h]). (iv) In the structure -N $^{\tau}$ V-.

In the recorded material, there are stretches identified as words with final consonantal portions [s t n]; this list may not be exhaustive, since in theory, [l r] could also occur word-finally.[11] Therefore no conclusive statement about the overall system of syllable (or word) final terms is made here.

---

[11] Finnish dictionaries list items such as *askel*, 'step', *manner* 'mainland'.

Exponents of τ

The exponent of τ is the apical articulation of the exponent of the word-final C-term.

Examples

55. jokaˑ $^{\zeta}$ õŋ ^{n} kĭn $^{\tau}$ hyʊ̃în $^{\tau ?}$ ĭlõînẽn ^{n} tauaˑtĕs:a:n $^{\tau}$
C—V $^{\zeta}$ V—N ^{n} P—N $^{\tau}$ C—N $^{\tau ?}$ V—N ^{n} P—N $^{\tau}$
{rel. pron.+nom. sg be+3ps+clitic very happy+nom meet+inf+iness+3pers. poss}
who is also very happy to meet (when she meets)

56. $^{?}$?ulos $^{\tau}$ tapa:mɑ̱:n $^{\tau ?}$?ɣst$^{?}$ɑ̱ʊɑ̰̃:^{n}sa ^{g}
$^{?}$ V—C $^{\tau}$ C—N $^{\tau ?}$ V-$^{?}$-V ^{g}
{out meet+inf+ill friend+part+3pers poss}
out to meet her friend

57. ɦaneŋ ^{n} kaʊel:ĕs:ɑ̰̃:n ^{n} talostŏ $^{\zeta}$ pois $^{\tau}$ pǎin $^{\tau}$
C—N ^{n} P—N ^{n} P—V $^{\zeta}$ C—C $^{\tau}$ C—N $^{\tau}$
{3ps+gen walk+inf+iness+3pers poss house+elat away direction}
as she walks away from the house

3.3 ?

Distribution of ?

? is found in two main structures: (i) when the second of two words is V-initial and the two words are not in what might be loosely called 'close grammatical contact' (see under ζ, Section 6.5.1.5 below), ie. in structures -C $^{?}$ V- and -V $^{?}$ V-; (ii) word-internally, where it frequently seems to be associated with resonant portions of long duration, such as long voiced lateral approximant portions, diphthongs or long vowels.

It should be pointed out that Itkonen (1965) shows that this type of word join is common only in the Savo dialects; and therefore the statement presented here, while accounting for my informants' speech, may not apply more generally in Finnish.

Exponents of ʔ

The exponents of ʔ include creaky voice. Creaky voice is timed in interesting ways with other phonetic parameters. Usually, the creaky voice coincides with changes in the vocal tract, so that any vowel transitions at the join between two words are, so to say, 'covered' by the creaky voice. This is the most usual pattern in stretches which expone **-V ʔ V-** structures. In stretches which are the exponents of **-N ʔ V-** structures, where the exponents of **-N** include nasality, the creaky voice is generally timed to coincide with the closing of the velum and the ending of nasal airflow. It may however also be timed so that a small amount of creaky voice and nasal airflow overlap; but when the creak comes to an end, nasality is not present.

Another feature of periods of creaky voice is that they often mark areas where the pitch changes. It is not uncommon to find creaky voice between a stretch that ends with a low pitch, and followed by one which begins with a high pitch.

For reasons which remain unclear, diphthongs and long vocalic or resonant consonantal portions are all susceptible to creaky voice. In the case of diphthongs, the creak tends to start at the end of the steady state portion of the initial part of a diphthong. Otherwise, creak is timed to start coincidental with the onset of the resonant portion. It may be true to say that creaky voice is a sort of masking technique: a way to cover up transitions from one state to another. It remains unclear what function (if any) creaky voice may have word-internally. It could be that there is just a conventional phonetic association in Finnish between resonant articulations, the exponents of length, and creaky voice.

The duration of creaky voice is anything between 20 and 160 ms. These are extremes, however. It is most usual in the material collected to find creaky voice with a duration of approximately 60 ms (±20 ms). Sometimes the glottal constriction is so tight as produce periods of complete glottal closure; these are generally released into creaky voice. Therefore, it would be inaccurate to describe these portions as 'long glottal stops' (cf. Itkonen 1965). Portions such as these are generally associated with creaky voice of greater duration.

Examples:

58. nɑinen τʔ i̱s·tu:
C—N τʔ V—V
{woman+nom sit+3ps}
a woman is sitting

59. jal:e:n ⁿ tɑkɑ· τʔ ŋa a:řes:ăʰ h
C—N ⁿ P—N τʔ V—V h
{again fireplace+gen edge+iness}
back by the fire again

60. h x̞ɑunis τʔ u̱:sĭ ζ mɑt:ð
h C—C τʔ V—V ζ C—V
{beautiful+nom new+nom rug+nom}
a lovely new rug

61. ʔ ʔulos τ tɑpɑ:mɑ̱:n τʔ ʔy̱stʔɑ̱ʊã̱:ⁿsa g
ʔ V—C τ C—N τʔ V-ʔ-V g
{out meet+inf+ill friend+part+3pers poss}
out to meet her friend

62. purkʔɑu̱tuɑ
C-ʔ-V
{come undone+inf}
to come undone

3.4 **g**

Distribution of g

g occurs at the junction of certain morphological items with other words. Itkonen (1965) lists nine structural places where **g** occurs, of which the most important are: negative present tense forms; 2ps imperatives; first infinitive; most nouns which end in [-e]; the third person personal suffix (singular and plural), which has the phonetic exponents [nsɑ, nsa] and adverbs marked with the suffix whose exponents are [sti]. In all these cases, **g** is a property of the end of the named elements of structure. The vast majority of Finnish words that end in [e] are joined to the next word with **g**.

In the data collected, there are relatively few instances of structures where **g** applies. There are one or two instances of negatives, and a few instances of 3rd person personal suffixes with the exponents [nsɑ, nsa].

It seems reasonable from the available data to conclude that **g** only occurs in structures with the general shape **-V ^{g} C-**, where **C-** stands for a C-term whose exponents include oral stricture. Most studies of 'gemination' in Finnish include the possibility of the structure **-V ^{g} V-**, but the cases of this in my data have exponents which are not distinguishable from the exponents of the structure -V $^{?}$ V-; since it simplifies the statement of exponents and is within the terms of the Principle of Reusability, I treat all the examples of potential -V ^{g} V- as the structure -V $^{?}$ V-.

Exponents of g

The exponents of g include the prolonged duration of the closure phase for the succeeding consonant, where 'closure' means any consonantal stricture. Articulations which could be described as more tense are also frequently found as exponents of **g** pieces. For instance, short [ʋ], a labiodental approximant, is found as the exponent of a C-term which only occurs initially in the syllable; but the same C-term in conjunction with **g** may have the exponent [v:], with a closer stricture as well as greater duration. Plosive bursts in **g** pieces are also frequently sharper than in non-**g** pieces.

Examples

63. (pp all ɦaŋ̊čŋ̊. ^{n} all pp) kis:ɑ̈nsð ^{g} k:ɑt·so:
C—N ^{n} P—V ^{g} P—V
{3ps+gen cat+3pers. poss look+3ps}
her cat his watching

64. nɑinẽ^{n} ^{n} sɑ: $^{\varsigma}$ li:nɑnsð ^{g} v:ɑlmi:ksi
C—N ^{n} C—V $^{\varsigma}$ C—V ^{g} C—V
{woman+nom get+3ps scarf+3pers. poss ready+transl}
the woman finishes her scarf

65. mut:ɑ $^{\varsigma}$ ɦan$^{\tau}$ ei $^{\varsigma}$ ɦuomɑ: ^{g} k:ɑ:n $^{\tau}$ et:ɑ̠̆ $^{?}$
C—V $^{\varsigma}$ C—N $^{\tau}$ V—V $^{\varsigma}$ C—V ^{g} C—N $^{\tau}$ V—V $^{?}$
{but 3ps not+3ps notice+pres emphatic clitic comp}
but she doesn't even notice that

66. ʋi̤:doin $^{\tau}$ ɦe ^{g} p:a:s̈ˤʋat $^{\tau}$ koti· $^{\varsigma}$ oʋe l:ẽ^{h} ^{h}
C—N $^{\tau}$ C—V ^{g} C—C $^{\tau}$ C—V $^{\varsigma}$ V—V ^{h}
{finally 3ppl arrive+3ppl home+door+all}
finally they get to the front door

Descriptions of Finnish phonetics (eg. Itkonen 1965) frequently describe long glottal plosives as the exponent of the join between two words where one ends in a vowel and the next starts with a vowel, and where the first word is joined to consonant-initial words with greater duration of the initial consonant. This would lead us in the terms of the present analysis to posit the structure **-V $^{g?}$ V-** to complement the structure **-V ^{g} C-**. Greater duration would be allotted as the exponent of **g**, and the glottal stricture as the exponent of **?**. However, in the few cases in the material where such a structure might apply, it seems not to. The phonetics of such potential structures is indistinguishable from the phonetics of the structure **-V $^{?}$ V-**, and therefore I have chosen to state the distribution of **g** in terms of the structure **-V g C-** only. For example, in the stretch

67. g t:ɑk:ɑ̆nsɑ ʔ ʔa̲:res:a
g P—V gʔ V—V
{fireplace+3pers poss edge+iness}
by her fire

the stretch of creaky voice lasts approximately 85ms. We may expect to find the exponents of **g** in this stretch of phonetics, since we find greater duration in other places where the third person possessive suffix precedes another word. However, in the stretch

68. kɑulɑ ζ li:nɑʔ˞ ʔ ʔɑlkɑ:
C—V ζ C—V ʔ V—V
{neck+scarf+nom start+3ps}
the scarf starts

the period of creaky voice lasts approximately 160ms. This is almost twice as long as the duration of the stretch of glottal constriction in the example which potentially has **g**, but this is counterintuitive. The long duration could also not justifiably be said to be the exponent of **g**, since **g** is not otherwise used to put together the noun *liina* with some other word, nor any other pair of words, except where the first one ends in [-e]. It may also be fair to say that the material collected here is so small that no firm conclusions can be drawn from it.

3.5 ζ

Distribution of ζ

ζ occurs in all cases where the structure of the junction is **-V C-**. This is the commonest junction in Finnish, since most words end with **V** and most words begin with **C** (Wiik 1977). The most commonly found inter-word structure is **-V ζ C-**.

ζ is also found in those **-V V-** structures to which **ʔ** does not apply: between words which are in what we might characterise as 'close contact'. This includes junctions with function words such as *mutta*, but, *ja*, and; the combination of *sanoa*, to say, *+ettă*, the complementiser; the negative verb; the verb *olla*, to be; and also between two items in a compound word where the first of them is V-final, and the second is V-initial. There is also a case in the data where ζ is found between a verb and the reflexive *itse*.

Exponents of ζ

The exponents of ζ include the presence of an open vocal tract accompanied by voicing followed by either a consonantal stricture with the same resonance as the subsequent part of the word or a vocalic portion, in which case the junction between the two vowels is marked by the absence of any glottal constriction, which is one exponent of ʔ. A change in resonance between front and back or back and front is one possible exponent of ζ, but is not criterial of ζ at word junctions.

Examples

69. ʋierestă ζ jă ζ lamːitːeleː ζ tɑkɑ̠ˑnɑ̯̆^{h} ^{h}
C—V ζ C—V̯ ζ C—V ζ C—V ^{h}
{side+elat and warm+3ps behind+ess}
…from the side and warms itself behind…

70. ^{ʔ} ʔystaʋalːeːn ^{τʔ} ųːtːɑ ζ ɦienoɑ ζ kɑulɑliːnɑː^{h} ^{h}
^{ʔ} V—N ^{τʔ} V—V ζ C—V ζ C—V ^{h}
{friend+all+3pers. poss new+part fine+part neck+scarf+part}
(to) her friend the fine new scarf

71. mutːɑ̆ ζ ystaʋɑ̳ ζ ɦuomɑː ζ kin ^{τ}
C—V ζ V—V ζ C—V ζ C—N ^{τ}
{but friend+nom notice+3ps clitic}
but the friend notices as well

72. jɑ ζ ɑlkɑː ζ neuloɑ̤ ^{h}
C—V ζ V—V ζ C—V ^{h}

{and start+3ps knit+inf}
and starts knitting

73. koti· ^{ζ} oʋel:ĕ^{h} ^{h}
C—V ^{ζ} V—V ^{h}
{home+door+all}
to the front door

3.6 h

Distribution of h

h is found finally and sometimes initially in the utterance. It marks initiality and finality. Not all initials nor finals are marked with **h**.

Exponents of h

The exponents of **h** remain somewhat inconclusive. They involve absence of regular vocal fold vibration (ie. presence of breathy voice, creaky voice, whispery voice, or simply release of air through the vocal tract). They may also involve relative more open, laxer, articulations. They may also involve the aspiration of plosives, and even slight affrication.

Examples

74. jɑ ^{ζ} nɑinĕŋ ^{n} keit:a: $^{?}$ a̲y̲ ystaʋa$^{?}$l̠:ɛŋ ^{n} kɑĥʋi̠t^{h} ^{τh}
C—V ^{ζ} C—N ^{n} P—V $^{?}$ V-$^{?}$-N ^{n} P—C ^{τh}
{and woman+nom cook+3ps friend+all+3pers. pos coffee+acc. pl}
and the woman makes her friend coffee

75. ^{h} sɕ̤ ^{ζ} o$^{?}$n̠: ^{τ} hỹʊ̃in ^{n} ty:pil:i̠sta̠h ^{h}
h C—V z V-/-N ι C—N n P—V h
3ps be+3ps very typical+part. sg.
it's quite typical

76. ^{h} x̯aunis $^{t?}$ u̯:sĭ ζ mat:ð

h C—C t/ V—V z C—V
{lovely+nom new+nom rug+nom}
a lovely new rug

77. sulke: ζ ʋerfiot^{S} th

C—V z C—C th
{shut+3ps curtain+nom.pl}
closes the curtains

3.7 The verb *olla*, to be

For the structure -C/V $^{\pi}$ V-, the usual term of π is ?. However, when words are in what I loosely termed 'close grammatical contact', they are more frequently joined by ζ. In this section, I shall consider in more detail the phonetics of the verb *olla*, to be, which exhibits rather complex word joins. This shows that the analysis presented in 3.1-6 is partial, and points to the need for an even more refined statement than the one given in this paper.

Examples 78-80 show the verb *olla* linked with ζ:

78. jă zě on̪
ja se on
and it is

79. ni: $_{\{all}$ ma ðm $_{all\}}$ pi:rtanyh
niin, minä (mä) olen (oon) piirtäny(t)
yes, I've drawn (it)

80. ei ne koʋĭ ?iso:ją ǫ: ną: kųką·t
ei ne kovi(n) isoja ole (olo) nää kukat
they're not very big, these flowers

There are in fact a variety of ways in which the verb *olla* or its parts may be joined to the preceding items. One of the common frames in my data is 'these are —'. For this, the Standard form is *nämä ovat*.. My informants' productions typically resemble those at (81).

81(a) {$_p$ nãm^ʏ·ɑ̯·t $_p$}
81(b) {$_p$ nãmaɑ̯·t $_p$}

It can be seen that the initial part is always [nãm-]. Then there is an open portion which has some labiality in it and is dark, though the darkness may vary in its domain from the nasal portion to the end, or not start till later in the second syllabic portion. It is difficult to know how many syllables there in in these utterances; but it is certainly not the four implied by the orthography. For the phrase 'they are unemployed' my informants produced:

82. {$_p$ he^{w}ɐ̯ $_p$}t:^{ʏ}øt:ømijḁ̆
he ovat työttömiä
they are unemployed

where it can be seen that there is labiality, but the expected amount of syllabicity is not present. A more extreme form of this lack of syllabicity as a distinct exponent of the verb *olla* can be seen in examples such as:

83. ɑɦmătil:ɑ̆·n ʔiso· mɑɦɑ·
ahmatilla on iso maha
the greedy person has a big stomach

84. ketɑ̃·ŋ kolɔ̃:m pu:ʈɑr̥hɑ̥̆s:ɑ̥̆
ketun kolo on puutarhassa
the fox's den is in the garden

In these cases, greater duration of the word-final vowels of the items just before the verb followed by nasality seems to be doing the work of the third person singular form of *olla*. In many instances, then, the verb *olla* seems to behave almost as if it were a clitic, and forms a special piece with the preceding item in the sequence of the speech. Much of the phonetics typical of other items with apparently similar phonological structure (i.e. -V V- pieces) is not to be found, and much of the phonetics of this verb is unlike that which is to be found with other verbs.

Frames such as *nämä ovat* and pieces where the items before the verb *olla* end in anything other than complete oral closure are commonly marked as 'lax' in my records: they tend to be articulated quickly, with less close stricture, more breathiness, and with unclearly differentiated syllables (i.e. it is often hard to say how many syllables one hears). They are often also quieter. Perhaps surprisingly, when the item before the verb ends in a consonant with complete oral stricture (with or without nasality as well), this portion of complete closure can be long before the verb *olla*:

85. han: on t^{Y}øtøn

hän on työtön
s/he is unemployed

86. ɑh̃mɑtit: oʋɑt keit:jøs:a

ahmatit ovat keittiössä
the greedy people are in the kitchen

In these cases, the way in which the word before the verb *olla* and the verb itself are joined phonetically is different from what is described above. Rather than having a juncture where material seems to go missing, here the juncture seems to be marked by 'more' material, i.e. greater duration. This could be treated as an exponent of **g**; however, it is the final consonant of the first item which is long, whereas in other cases where **g** joins words, such as the imperative, it is the initial consonant of the second item which is long.

Itkonen (1965: 248-265) discusses both these kinds of word join across the Savo dialect area, and notes that in his data most examples of -C$^{\zeta}$ V- (cf. exx. 78-80) involve the verb *olla* and the negative verb *ei*. Itkonen observes that this junction can only occur with 'close-knit compounds'. He also notes the junctions with long consonantal portions, and claims that they contain two distinct intensity peaks, something which I did not observe with my informants. They are also rare in his material. While no clear conclusions can be drawn, it does seem clear that not all items can be handled in the same way in any complete analysis of Finnish word joins.

3.8 Spectrograms of examples of inter-word junctions

Figures 7-10 below show spectrograms of some of the utterances described in the previous section. The relevant details are commented on in conjunction with the appropriate spectrogram. The spectrograms are provided to show that phonetic exponency can be made to account to more than one kind of phonetic description.

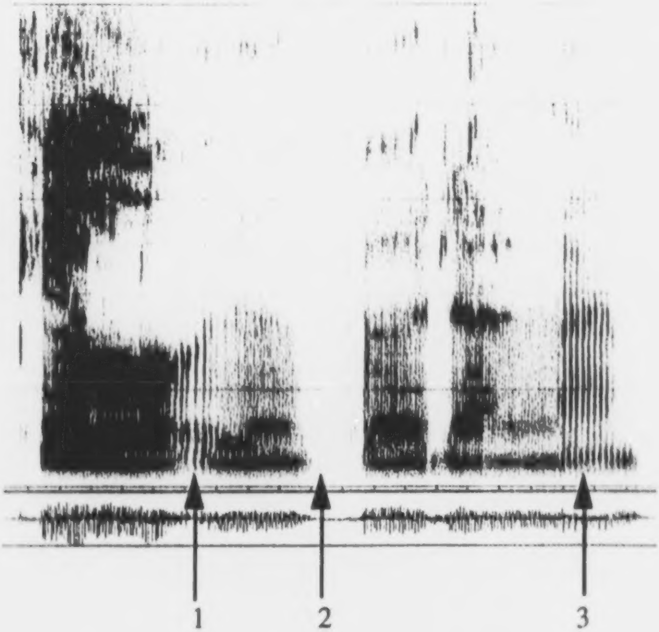


Fig.7: Spectrogram of 'Nainen keittää ystävälleen kahvit'.

Note that the temporal extent of voicing between the nasal and plosive portions is different at (1) and (3) in Fig. 7 above; this provides good evidence that temporal information is properly part of the phonetic exponency. The period of creaky phonation around (2) lasts approximately 130ms; this is approximately twice as long as other stretches of creaky voice in Figs. 34-36, yet there is no motivation for saying that the duration of this portion of creak is an exponent of **gʔ** rather than just **ʔ**. Note that the final plosive burst is rather diffuse, aspirated, and does not have such a well-defined burst as at (1) and (3); this lax articulation is an exponent of **h**. The structure of the whole utterance, then, is **C—N**^{n} **P—V**^{ʔ} **V—N**^{n} **P—C**^{h}.

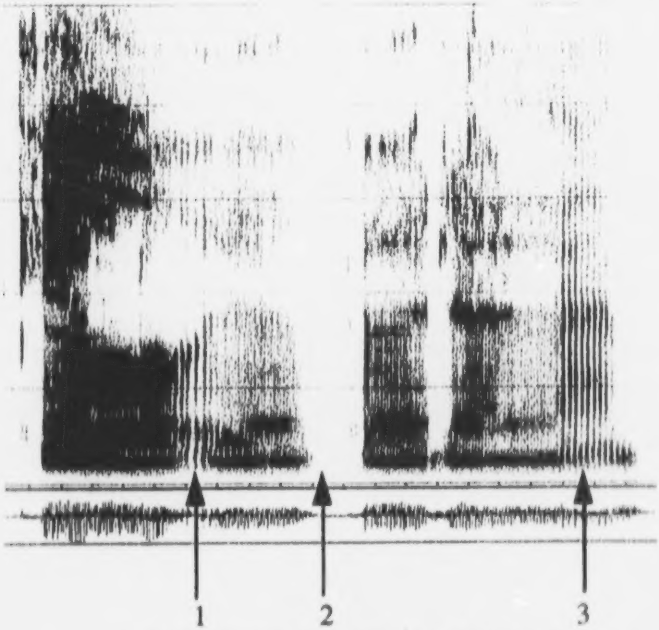


Fig. 8: Spectrogram of 'Kello yöpöydällään'.

In Fig. 8, note the creaky voice at (1), which extends for about 60ms. Note also that the formant transitions are timed to coincide with this stretch of creak, so that the non-creaky portions before and afterwards contain more or less steady state formants. At (2) are the exponents of ζ, a voiced vocalic portion followed by a portion with consonantal stricture. Note how at (3) the creaky voice is timed to coincide exactly with the release of lateral airflow, thus masking any formant transitions out of the lateral. It remains unclear why creaky voice should associate with stretches such as long vowels. The phonological structure for this utterance is **C—V ʔ V—V ζ C-ʔ-N**, since the word *yöpöytä* is a compound noun, *yö* 'night' + *pöytä*, 'table'.

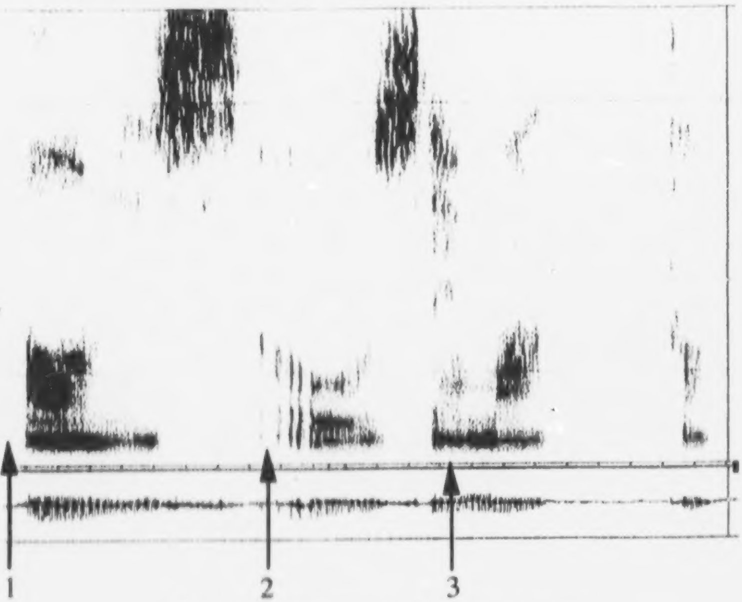


Fig. 9: Spectrogram of 'Kaunis uusi matto'.

Fig. 9 shows the spectrogram for *kaunis uusi matto*, 'a lovely new carpet'. In this case, attention is drawn to the lax articulation of the initial voiceless portion, which has a sudden onset, but lacks a clearly-defined burst, at (1); this is taken to be an exponent of **h**. Note that at (2) the exponents of **ʔ** are evident, and that the creaky voice is timed to coincide with the transitions from the preceding consonantal constriction into the vocalic portion at the beginning of the second word. At (3) the exponents of **ζ** are again evident from the unmarked transition from the vocalic portion at the end of one word and the consonantal portion at the start of the next. The structure of this phrase is **h C—C τʔ V—V ζ C—V**.

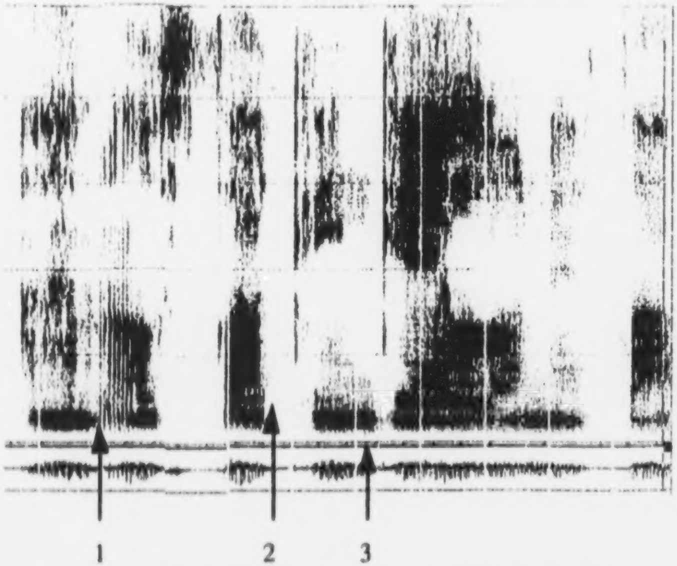


Fig. 10: Spectrogram of 'Hän ostaakin kellon jota...'

Fig. 10 shows the spectrogram of the phrase *hän ostaakin kellon jota...* 'and he buys the clock which...'. Note at (1) the exponents of **ʔ**; in this case the creak lasts for about 50ms. Note how again the creaky voice is timed to coincide with the offset of the consonantal articulation and thus covers the portion of the acoustic signal which exhibits the greatest amount of formant transitions. The portions at (2) and (3) can be usefully compared, since both show velar closure followed by a plosive release. At (2) the closure is clearly unvoiced, and the structure is —V ζ P—, since the two words are in close grammatical contact (verb + clitic). At (3) on the other hand, there is obvious voicing in the closure portion; this is attributable as an exponent of **n**. The overall structure of the phrase then is C—N ʔ V—V ζ P—N n P—C τ C—V.

3.9 Summary

Tables 1 and 2 present (i) the structures found in inter-word position, and (ii) the statement of exponents in broad terms of the inter-word prosodies.

Word-Final	Inter-word Prosody	Word-Initial
-N	n	C- (C-' = [p t k m n s l r ʋ])
-C or -V	ζ when in close grammatical contact; ʔ otherwise	V-
-C	τ	C-, V-, or utterance final
-V	g when morphology demands it; ζ otherwise	C-
-C or -V	h	utterance-final
utterance-initial	h	C- or V-

Table 1: Summary of the inter-word structures.

More than one statement above may apply, and two prosodies of inter-word junction may be combined; the structures -C$^{\tau\,ʔ}$ V- and -C$^{\tau h}$ # are possible, and do not contradict the above statements.

n	sameness of place of articulation of exponents of -N and C-.
ʔ	creaky voice timed to coincide with changes in the vocal tract.
ζ	vocalic articulation followed either by a consonanatal articulation (in $^{\pi}$C- structures) or by a vocalic articulation with no intervening glottal constriction (in $^{\pi}$V- structures).
h	voicelessness, creaky voice, breathy voice or exhalation; laxer and more open consonantal articulations.
τ	apical articulation of -C.
g	long duration of C-.

Table 2: Summary of the broad exponents of the inter-word prosodies.

4. Conclusion

This paper has shown how a phonological statement can be made which takes into consideration phonetic characteristics which in most phonologies are considered irrelevant. Some of its important characteristics are:

1. A parametric phonetic statement is made in either acoustic or articulatory phonetic terms.
2. The phonological statement is made in phonological terms, which are abstract in the sense that they have no implicit phonetics.
3. The two levels of phonetics and phonology are connected by statements of phonetic exponency. These exponency statements need not be simple, in the sense that they may refer to more than one phonetic parameter (cf. Ogden 1995a).
4. The exponency statements account for what might be characterised as 'fine phonetic detail'. The resulting analysis is therefore based on, and accountable to, observed phonetic detail, some of which would be deemed irrelevant if an analysis were used which were based on

a phoneme concept, or which could only produce a broad phonetic level of description, such as most current work in generative phonology.

5. The phonological statement presented describes in declarative, non-process terms features of Finnish which are otherwise typically regarded as processes of assimilation, or the output of a series of rules; or ignored altogether.

6. The phonological statement makes reference to other levels of linguistic statement such as the morphosyntactic and interactional levels. Thus there is integration of different levels of lingusitic statement.

REFERENCES

Catford, J. C. (1988). *A Practical Introduction to Phonetics.* Oxford: Clarendon Press.

Docherty, G, Foulkes, P., Milroy, J., Milroy, L. (1995). *What is a phonlogical fact? The relationship between theory and data.* Newcastle University, ms.

Engstrand, O & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: cross-language evidence for a theory of adaptive dispersion. *Phonetica* 51. 80-91.

Flifilet, A. L. (1971). Une évaluation structurale du système syllabique finnois. In L.L. Hamerich, Roman Jakobson, Eberhard Zimmer (eds.). *Form and Substance.* Odense: Akademisk Forlag. 193-203.

Hawkins, S. & Slater, A. (1994). Spread of CV and V-to-V coarticulation in British English: implications for the intelligibility of synthetic speech. *Proc. ICSLP-94* 1. 57-60.

Hawkins, S (1995). Arguments for a nonsegmental view of speech perception. *Proc ICPhS* 3.18-25.

Henderson, E.J.A. (1971). *The Indispensable Foundation. A Selection from the writings of Henry Sweet.* London: Oxford University Press.

Itkonen, T. (1965). *Proto-Finnic final consonants. Their history in the Finnic Languages with particular reference to Finnish Dialects.* Helsinki: SKS.

Karlsson, F, (1971). Finskans rotmorfemstruktur: en generativ beskrivning. *Turun Yliopiston Fonetiikan Laitoksen Julkaisuja* 10. [The structure of Finnish stems: a generative description.]

Karlsson, F. (1982). *Suomen kielen äänne- ja muotorakenne*. Juva: WSOY. [The phonology and morphology of the Finnish language.]

Kettunen, L. (1981). *Suomen murteet: Murrekartasto*. Helsinki: SKS. [Finnish dialects: dialect atlas.]

Lahti, L.-L. (1981). On Finnish plosives. *Lund working papers in Linguistics and Phonetics* 21. 89-93.

Lehiste, I. (1965). Juncture. *Proceedings of the Fifth International Congress of the Phonetic Sciences*, Basel/New York: S. Karger, 172-200.

Local, J & Ogden, R. A. (1994). A model of timing for non-segmental phonological structure. *Proceedings of ESCA/IEEE workshop on Speech Synthesis*. 236-239.

Manuel, S. Y, Shattuck-Hufnagel, S., Huffman, M., Stevens, K.N., Carlsson, R. & Hunnicutt, R. (1992). Studies of vowel and consonant reduction. In *Proceedings of the Internaitonal Conference on Speech and Language Processing*. Vol. 2. 943-946.

Ogden, R. A. (1995a). *An exploration of phonetic exponency in Firthian Prosodic Analysis: form and substance in Finnish phonology*. D.Phil., University of York.

Ogden, R. A. (1995b). Where is timing? A response to Caroline Smith. Paper presented at LabPhon 4, Oxford. To appear in Amalia Arvaniti & Bruce Connell (eds) *Papers in Laboratory Phonology 4*. Cambridge: CUP.

Palomaa, J. K. (1946). Suomen kielen äännekestoista puhumaan oppineen kuuromykän ja kuulevan henkilön ääntämisessä. Publicationes Intituti Phonetici Universitatis Helsingiensis. [On the durations of Finnish sounds in the pronunciation of a deaf mute who has been taught to speak and a hearing person.]

Rapola, M. (1966). *Suomen kielen äännehistorian luennot*. Helsinki: SKS. [Lectures on the history of the Finnish language.]

Sovijärvi, A. (1938). *Die gehaltenen, geflüsterten und gesungenen Vokale der finnischen Sprache*. Helsinki: SKS.

Sovijärvi, A. (1957). The Finno-Ugrian Languages. In L. Kaiser: *A Manual of Phonetics*. Amsterdam: North-Holland. 312-324.

Sprigg, R. K. (1957). Junction in spoken Burmese. *SLA*. 104-138.

Suomi , K. (1980). Voicing in English and Finnish stops. A typological comparison with an interlanguage study of the two languages in contact. *Turun yliopiston suomen laitoksen julkaisuja* 10.

Suomi, K. (1985a). Yleissuomen /i/—/j/ ja /u/—/v/-oppositioiden rakenteellisista perusteluista. *Virittäjä* 89, 48-56. [On the structural motivations of the Finnish /i/—/j/ and /u/—/v/ oppositions.]

Whitley, E. (ms). Phonetic Analysis and Phonological Statement. Notes from a lecture course entitled *Phonetic Analysis and Phonological Statement*, taken by Eugénie Henderson. York: Firthian Prosodic Archive.

Wiik, K. & Lehiste, I. (1968). Vowel quantity in Finnish Dissyllabic words. *Congressus Secundus Internationalis Fenno-Ugristrarum*, I, Helsinki: Suomalais-ugrilainen Seura. 569-574.

Wiik, K. (1965). Finnish and English Vowels. *Turun Yliopiston fonetiikan laitoksen julkaisuja* 94. Turku.

Wiik, K. (1975). On vowel duration in Finnish dialects. *Congressus Tertius Internationalis Fenno-Ugristrarum*, I. Tallinn: Valgus. 415-424.

Wiik, K. (1977). Suomen tavuista. *Virittäjä* 81, 265-278. [On Finnish syllables.]

Wiik, K. (1981). *Fonetiikan perusteet*. Juva: WSOY. [The foundations of phonetics.]

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in stop sequences. *Journal of Phonetics* 22. 121-140.

OLD ENGLISH VERB-COMPLEMENT WORD ORDER AND THE CHANGE FROM OV TO VO*

Susan Pintzuk

Department of Language and Linguistic Science
University of York

1. Introduction

The change from object-verb (OV) word order to verb-object (VO) word order is one of the most striking changes in the history of the English language. According to most generative accounts, Old English is an OV language, with optional rules of postposition and some form of the verb-second (V2) constraint. Modern English, of course, is a VO language and exhibits only remnants of V2.[1] The change from OV to VO is usually described as an abrupt grammatical reanalysis occurring at the end of the Old English period.[2]

This paper offers an alternative account of Old English verb-complement word order and the change from OV to VO. Evidence is provided that the change does not involve abrupt reanalysis but rather

---

* The original version of this paper was presented at the Eighth International Conference on English Historical Linguistics in Edinburgh, Scotland, 19-23 September 1994. Thanks are due to two anonymous reviewers for suggestions and comments. Author's e-mail: sp20@york.ac.uk.

[1] For example, Modern English shows residual V2 effects in questions and in clauses with preposed negative polarity items:

(i) What should I do?

(ii) Never have I seen such a sight.

[2] There are three stages in the history of English: Old English (700-1100), Middle English (1100-1500), and Modern English (1500-present).

York Papers in Linguistics 17 (1996) 241-264
© Susan Pintzuk

synchronic competition between two grammars, which begins in the Old English period and continues during the Middle English period.

The paper is organized as follows. Section 2 presents background assumptions and terminology. Section 3 describes in more detail the standard analysis of Old English and the change from OV to VO. Section 4 presents three predictions of the standard analysis and shows that they are not fulfilled. And Section 5 proposes an analysis of grammatical competition to account for the variation in verb-complement word order during the Old and Middle English periods.

The proposed analysis is based upon an investigation of data collected from sixteen Old English texts; for sampling techniques and information about the texts included in the database, see Appendix B of Pintzuk (1993). Old English texts are cited according to the system specified in Mitchell, Ball, and Cameron (1975, 1979); the abbreviations used are listed in the Appendix.

2. Background assumptions and terminology

The analyses presented in this paper use a generative approach to describe syntactic structure and word order, the Principles and Parameters framework outlined in Chomsky (1981, 1986) and related work. In particular, it is assumed that the base component of the grammar generates underlying structure and word order that are modified by syntactic movement, deriving surface structure and word order; both structure and movement are constrained by universal principles. The differences between languages, and between different stages of the same language, are described in terms of parameters; for example, one difference between Modern German and Modern English is the setting of the parameter that determines the order of verbs and their complements. For ease of exposition, I make the following three assumptions about the syntax of Old English: (i) there are only two functional categories, Infl and Comp; (ii) the underlying order of heads and their complements can vary; and (iii) only finite verbs move from their underlying position to functional heads. Nothing crucial rests on these assumptions or on the choice of this particular framework: the syntactic differences between OV and VO languages and grammars are robust and can be expressed in any framework.

The term 'auxiliary verb' is used for expository convenience to refer to those verbs that take infinitival or participial complements in Old English.[3] The terms 'verb raising' and 'verb projection raising' are used to describe the permutation of auxiliary verbs and their infinitival or participial complements in otherwise verb-final languages.[4] The term 'heavy constituent' is used for Old English PPs, non-pronominal NPs, polysyllabic adverbs, and non-finite verbs, to distinguish them from 'light constituents', i.e. pronouns, particles, and monosyllabic adverbs.[5] The terms 'OV' and 'VO' are used to refer to either underlying or surface word order and structure; the use will be made clear by the context. The term 'Infl-medial' is used for structures where Infl, the head of IP, precedes its complement; the term 'Infl-final' is used for structures where Infl follows its complement.

It is assumed that Old English is a V2 language, although the precise formulation of the V2 constraint for Old English is still a matter of some debate (see, for example, van Kemenade 1994, Pintzuk 1993); and that finite verbs obligatorily move to Infl to receive inflection. Because leftward verb movement to a functional head can distort the underlying word order in both main and subordinate clauses, it is necessary to abstract away from this effect in order to focus upon the order of verbs and their complements. The structural ambiguity is illustrated below: clauses like (1a), with the finite main verb in clause-medial position, can be derived either by leftward movement of the verb, as in (1b), or by rightward movement of the post-verbal constituent, as in (1c).

---

[3] Allen 1975 shows that Old English does not have a separate word class of auxiliary verbs. But see Warner 1993 for features of a subset of my Old English auxiliaries that distinguish them from lexical verbs.

[4] See den Besten and Edmondson 1983, Evers 1975, 1981, Haegeman 1994, Haegeman and van Riemsdijk 1986, Kroch and Santorini 1991, among others, for formal analyses of verb (projection) raising in Germanic languages. No position is taken here on the derived structures of verb raising and verb projection raising. These processes are grouped with postposition in Section 3 simply on the basis of derived word order.

[5] It is shown in Pintzuk 1994 that Old English pronouns and adverbs behave differently from heavy constituents: they can be syntactic clitics, moving leftward to attach to maximal projections and/or heads.

(1) a. þe god worhte þurh hine
which God wrought through him
'... which God wrought through him ...'
(ÆLS 31.7)

b. Leftward verb movement:
þe god worhte$_i$ þurh hine t$_i$

c. Rightward movement of the PP:
þe god t$_i$ worhte [$_{PP}$ þurh hine]$_i$

To avoid this ambiguity, the data that will be considered here consist mainly of clauses with finite auxiliary verbs and non-finite main verbs; in these clauses the position of the auxiliary verb may be affected by V2, but the non-finite main verb remains in its base-generated position.[6]

3. The standard analysis of Old English

In this section the standard analysis of Old English, as proposed or assumed by van Kemenade (1987), Koopman (1990), Lightfoot (1991), and Stockwell and Minkova (1991), among others, is considered in more detail. According to this analysis, Old English has underlying OV structure, some form of V2, and postposition rules moving various constituents rightward beyond the main verb of the clause. All surface word orders are derived from a uniform base by optional movement rules, as illustrated in the examples below.[7] In (2), the underlying and surface order of the main verb and its complement are the same; in (3), VO surface word order is derived from OV underlying word order by postposition of the NP.

---

[6] Higgins 1991 suggests that Old English infinitives may move to the Infl position of the embedded non-finite clause; see Pintzuk 1991 for criticism of this analysis.

[7] Since the focus of this paper is the order of main verbs and their complements, the traces of topics and verbs affected by V2 are not shown in the examples.

(2) OV surface word order:

he ne mæg his agene aberan
he not may his own support
'He may not support his own.'
(CP 52.2)

(3) VO surface word order:

þu hafast t_i gecoren $[_{NP}$ þone wer$]_i$
you have chosen the man
'You have chosen the man.'
(ApT 23.1)

There is strong evidence in favor of this analysis, which forms the basis of most of the current work in Old English syntax within a Principles and Parameters framework. Evidence for underlying OV word order is provided by clauses in which main verbs follow their complements and auxiliary verbs follow the main verbs, as in (4). Evidence for the postposition of NPs and PPs and for verb (projection) raising is provided by clauses in which the finite auxiliary is preceded by two or more heavy constituents and followed by an NP, as in (5), a PP, as in (6), a non-finite main verb, as in (7), or a projection of the non-finite main verb, as in (8). Note that none of the clauses in (4) through (8) can be analyzed as V2 clauses, since the finite auxiliary is preceded by more than one heavy constituent.

(4) Evidence for underlying OV word order:

him þær se gionga cyning þæs oferfæreldes forwiernan mehte
him there the young king the crossing prevent could
'... the young king could prevent him from crossing there.'
(Or 44.19-20)

(5) Evidence for NP postposition:

þæt ænig mon t_i atellan mæge $[_{NP}$ ealne þone demm $]_i$
that any man relate can all the misery
'... that any man can relate all the misery ...'
(Or 52.6-7)

(6) Evidence for PP postposition:

her Cenwalh t_i adrifen wæs $[_{PP}$ from Pendan cyninge$]_i$
in-this-year Cenwalh driven-out was by Penda king
'In this year, Cenwalh was driven out by King Penda.'
(ChronA 26.19 (645))

(7) Evidence for verb raising:

Wilfrid eac swilce of breotan ealonde t_i wes $[_V$ onsend$]_i$
Wilfred also from Britain land was sent
'Wilfred was also sent from Britain.'
(Chad 162.27-164.28)

(8) Evidence for verb projection raising:

hwær ænegu þeod æt oþerre t_i mehte $[_{VP}$ frið begietan$]_i$
where any people from other might peace obtain
'... where any people might obtain peace from another ...'
(Or 31.14-15)

In anticipation of the discussion in Section 4.1, it should be pointed out that an OV grammar with optional rules of V2 and postposition is quite powerful and can derive many different surface word orders, some in more than one way. Because both leftward movement of the finite verb and rightward movement of NPs, PPs, verbs, and verb projections are permitted, the main verb can precede or follow its complement, and the auxiliary can precede or follow the main verb. This is illustrated in (9), where S = subject, XP = NP/PP complement, Aux = auxiliary verb, Vf = finite main verb, V = non-finite main verb.

(9)	Surface word order	Derivation
a.	S XP Vf	reflects underlying word order
b.	S Vf_i XP t_i	V2
c.	S t_i Vf XP_i	postposition
d.	S XP V Aux	reflects underlying word order
e.	S XP t_i Aux V_i	verb raising
f.	S Aux_i XP V t_i	V2
g.	S t_i Aux $[XP\ V]_i$	verb projection raising
h.	S t_i V Aux XP_i	postposition
i.	S Aux_i t_j V t_i XP_j	V2 + postposition
j.	S t_i t_j Aux V_j XP_i	verb raising + postposition

Given this analysis of Old English syntax, the following scenario is invoked to describe the change from OV to VO. During the Old English period, VO surface word order gradually increases in frequency at the expense of OV. Toward the end of the period, when the surface word order is overwhelmingly VO, language learners abduce a new grammar with underlying VO structure and word order on the basis of the VO primary linguistic data. During the transition period, when two grammatical systems are in use by the two different generations of speakers, clauses like (10a) are produced and understood under both the old and the new grammars, but with different analyses: under the old system, they are derived from OV structure by postposition, as shown in (10b); under the new system, they are derived from VO structure with no movement, as shown in (10c). One point deserves emphasis here. To the linguist, (10a) is structurally ambiguous and can be derived from one of two different underlying structures. But according to the abrupt reanalysis view of syntactic change, children abduce either the old OV grammar or the new VO grammar but not both, and the clause has a single underlying word order within each system.

(10) a. þu hafast gecoren þone wer
you have chosen the man
'You have chosen the man.'
(ApT 23.1)

b. Old OV grammar with postposition:
þu hafast t_i gecoren $[_{NP}$ þone wer$]_i$

c. New VO grammar:
þu hafast $[_{VP}$ gecoren þone wer$]$

The account presented above is both plausible and appealing. It depicts a period of word order variation generated by a uniform grammar, followed by the abrupt resetting of the parameter that controls the underlying order of verbs and their complements. And it offers an explanation for the change: the primary linguistic data used by children for language acquisition have changed, and therefore the grammar that is abduced differs in one or more parameter settings from the grammar of the previous generation. Despite its plausibility and appeal, however, it will be demonstrated in Section 4 that the predictions made by this analysis are not correct, and therefore that the analysis cannot be maintained.

4. Predictions of the standard analysis

The standard analysis of Old English and of the change from OV to VO presented above makes three predictions that can be tested on historical data. First, clauses unambiguously derived from the new VO grammar are not used during the Old English period, before the change. Second, clauses unambiguously derived from the old OV grammar are not used during the Middle English period, after the change. And third, the frequency of VO surface word order increases during the Old English period, to reach near categorical status in the primary linguistic data used by language learners. These three predictions are discussed in Sections 4.1 through 4.3.

4.1. Prediction #1: no VO clauses in Old English

According to the first prediction made by the standard analysis, we will not find Old English clauses that are unambiguously derived from the new VO grammar. Contra this prediction, it will be demonstrated below that clauses with underlying VO structure are used productively during the Old English period.

Although (9) above illustrates that an OV grammar with optional rules of V2 and postposition can derive many different surface word orders, there is one clause type that constitutes evidence for underlying VO word order. The relevant clauses are those with light constituents -- particles, pronominal objects, and monosyllabic adverbs. In Old English clauses with auxiliary verbs, these constituents appear both before and after the non-finite main verb, as shown in (11).

(11) a. Particle before the main verb:

and hi næfre siððan ut-brecan ne magon
and they never afterwards out-burst not may
'And afterwards they may never burst out ...'
(ÆCHom ii.174.3)

b. Particle before the main verb:

& woldon hig utdragan
and (they) would them out-drag
'... and they would drag them out.'
(ChronE 215.6 (1083))

c. Particle after the main verb:

he wolde adræfan ut anne æþeling
he would drive out a prince
'... he would drive out a prince ...'
(ChronB (T) 82.18-19 (755))

However, the position of these constituents varies only in clauses like (11b) and (11c), with the auxiliary verb before the main verb. In clauses like (11a), with the auxiliary verb after the main verb, particles, pronouns, and monosyllabic adverbs -- unlike heavier constituents -- invariably appear before rather than after the main verb. The distribution is shown in Table 1.[8]

Table 1
Distribution of particles, pronouns, and monosyllabic adverbs in Old English main clauses with auxiliary verbs

Clause Type	Before Main Verb		After Main Verb		Total
	N	%	N	%	
Main verb + aux	90	100.0%	0	0.0%	90
Aux + Main verb	260	94.5%	15	5.5%	275

It is obvious from the order of the main verb and the auxiliary that clauses like (11a) are OV in underlying structure, with the light constituent base-generated in pre-verbal position. The fact that light constituents never appear post-verbally in OV clauses indicates that these constituents cannot be postposed, probably because of a heaviness constraint on postposition. But if particles, pronouns, and monosyllabic adverbs do not postpose, then clauses like (11c) must be derived from underlying VO structure, as shown in (12); and these clauses therefore constitute evidence for the use of VO structure during the Old English period.

---

[8] The data for Table 1 consist of main clauses with particles from the database of Hiltunen 1983, supplemented by main clauses with pronominal objects and main clauses with monosyllabic adverbs.

(12) he wolde [$_{VP}$ adræfan ut anne æþeling]
he would drive out a prince

The position of the other constituents in the 15 clauses with post-verbal particles, pronouns, and monosyllabic adverbs lends further support to this analysis. In 14 of the 15 clauses, the auxiliary and main verb are adjacent, with all complements and adjuncts appearing after the main verb, as in (11c) above. The remaining clause, given in (13), has only an adverb between the auxiliary and the main verb.

(13) and man ne mihte swa ðeah macian hi healfe up
and one not could nevertheless put them half up
'... and nevertheless, one couldn't put half of them up.'
(ÆLS 21.434)

It must be concluded that the first prediction of the standard analysis is incorrect: VO structure is used productively, although perhaps at a low frequency, during the Old English period, before the change from OV to VO is supposed to have taken place.

4.2. Prediction #2: no OV clauses in Middle English

According to the second prediction made by the standard analysis, we will not find clauses in Middle English that are unambiguously derived from the old OV grammar. Contra this prediction, it will be demonstrated below that clauses with underlyingly OV structure are used productively during the Middle English period.

A number of studies demonstrate that OV surface word order, at least, is used in Middle English texts. Kroch and Taylor (1994) examine the position of NP complements in subordinate clauses with auxiliary verbs, where the order of the main verb and its complements is not affected by verb movement to Infl, in Early Middle English prose texts. In two West Midlands texts, they find a total of 23 out of 88 (26%) NPs in pre-verbal position between the auxiliary verb and the non-finite main verb; in three Southeast texts, they find a total of 31 out of 108 (29%) NPs in pre-verbal position. Stockwell and Minkova

(1991), citing Morohovskiy (1980), state that in 7.6% of the 14th to 16th century London texts, the complement appears before the main verb in clauses with auxiliary verbs. And Foster and van der Wurff (1993, 1994) show that OV surface word order is used productively throughout the Middle English period, although at a low frequency. Of course, we can't be sure how OV surface word order is derived in Middle English: it could reflect underlying structure and word order, as shown in (14a), or else be derived from a VO base by leftward movement, as shown in (14b).[9]

(14) a. Underlying OV structure:
S XP Vf

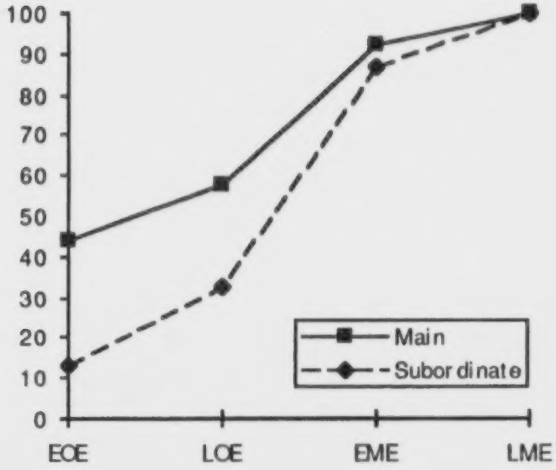
b. Underlying VO structure with leftward movement:
S XP_i Vf t_i

Clearly, the simple existence of clauses with OV surface word order is not sufficient evidence for OV underlying structure. But one clause type does provide evidence for OV structure in Middle English: clauses with pre-verbal particles. Since particles do not scramble leftward, pre-verbal particles directly reflect the underlying word order. As shown in Figure 1 (= Hiltunen 1983: 111, his Figure 2), particles appear before the main verb at a low but significant frequency throughout the Middle English period, in main clauses as well as in subordinate clauses, indicating that OV structure is used in Middle English.

---

[9] See Kroch and Taylor 1994 for speculations that the West Midlands dialect is mainly VO in underlying structure, while the Southeast dialect exhibits synchronic competition between OV and VO grammars.

Figure 1
Frequency of verb...particle word order in Early Old English (EOE), Late Old English (LOE), Early Middle English (EME), and Late Middle English (LME).



It is interesting to note that the discourse function of OV surface word order seems to be the same in Middle English as in Old English: Foster and van der Wurff (1994) demonstrate that pre-verbal position in Middle English is associated with inferable and evoked entities in Middle English; similarly, Linson (1993) shows that pre-verbal position in Old English is associated with entities that have been previously mentioned in the discourse.

It must be concluded that the second prediction of the standard analysis is incorrect: OV structure is used productively, although perhaps at a low frequency, throughout the Middle English period, after the change from OV to VO is supposed to have occurred.

4.3. Prediction #3: increase in VO surface word order

According to the standard analysis, the frequency of VO surface word order increased at the expense of OV surface word order during the Old English period, until it became nearly categorical. This section discusses the change in surface word order, the possible sources of the VO increase, and the role that the increase may have played in the change from OV to VO.

As a simple description of Old English word order, it is certainly true that VO surface word order was more common at the end of the period than in the earlier stages. Hiltunen (1983) shows that verb-particle word order was used more frequently in Late Old English than in Early Old English, both in main clauses and in subordinate clauses (see Figure 1 above); and Bean (1983) shows that OV word order decreased in frequency from the early to the late sections of the *Anglo-Saxon Chronicle*.

However, given the analyses presented above, there are at least four different ways to derive VO surface word order in Old English: (i) from OV structure, by leftward movement of the finite main verb, as in (15a); (ii) from OV structure, by postposition of the complement, as in (15b); (iii) from OV structure, by a combination of verb movement and postposition, as in (15c); and (iv) as a reflex of underlying VO structure, as in (15d) and (15e).

(15) a. Verb movement:
$S\ Vf_i\ [_{VP}\ XP\ t_i]$

b. Postposition:
$S\ [_{VP}\ t_i\ Vf\ ]\ XP_i$

c. Verb movement + postposition:
$S\ Aux_i\ [_{VP}\ t_j\ V\ t_i]\ XP_j$

d. Underlying VO structure:
$S\ [_{VP}\ Vf\ XP\ ]$

e. Underlying VO structure:
$S\ Aux\ [_{VP}\ V\ XP\ ]$

Researchers differ on the source of the increase in VO surface word order during the Old English period. Most scholars (e.g. Aitchison 1979, Canale 1978, van Kemenade 1987, Stockwell 1977) attribute it to an increase in the rate of postposition. Although the rate of postposition over time has not been measured for Old English, Santorini (1993) looked at the rates of NP and PP postposition in the history of Yiddish, a language that has undergone syntactic changes similar to English -- in particular, Yiddish changed from Infl-final to Infl-medial and from OV to VO. Santorini found that while the rate of postposition in structurally unambiguous clauses is highly variable from text to text, it does not increase over time. The data are shown in Table 2 below (= Santorini 1993: 275, Table 5). It is reasonable to conclude that the rate of postposition was not a factor in the OV to VO change in Yiddish, and it remains to be demonstrated that an increase in the rate of postposition played a role in the OV to VO change in the history of English.

Table 2

Rates of NP and PP postposing in Yiddish

	NP Postposing			PP Postposing		
Time period	Postposed	Not Postposed	Rate	Postposed	Not Postposed	Rate
1400-1489	1	12	8%	9	12	43%
1490-1539	7	19	27%	13	16	45%
1540-1589	7	24	23%	52	21	71%
1590-1639	10	40	20%	39	23	63%
1640-1689	4	19	17%	17	30	36%
1690-1739	1	5	17%	6	3	67%
1740-1789	1	2	33%	8	7	53%
1790-1839	0	1	0%	1	1	50%

In fact Lightfoot (1991) states that there is no evidence for an increase in the rate of postposition during the Old English period; he suggests instead that the source of the increase in VO surface word order in the primary linguistic data is an increase in the use of V2 in main clauses. Lightfoot shows that indicators of OV structure[10] are robust in languages like Dutch and German, but weak or non-existent in Old English. He suggests that an increase in VO surface word order derived by V2, coupled with the absence of evidence for OV structure, triggers the change from OV to VO.

In apparent support of Lightfoot's hypothesis, an increase in the frequency of clauses with the finite verb in second position is well documented: Pintzuk (1991), for example, demonstrates that for clauses with auxiliary verbs, the frequency of V2 in both main and subordinate clauses increases over the course of the Old English period.[11] But while V2 derives VO surface word order in clauses with finite main verbs and topicalized subjects, as in (16), it has no effect on the order of verbs and their complements in clauses with topicalized objects, as in (17), or in clauses with non-finite main verbs, as in (18).

(16) Philippus & Herodes todældun Lysiam
Philip and Herod divided Lycia
'Philip and Herod divided Lycia.'
(ChronA 6.4 (12))

(17) Of Iotum comon Cantware & Wihtware
From Jutes came people-of-Kent and people-of-Wight
'From the Jutes came the people of Kent and the people of Wight.'
(ChronA 12.13 (449))

---

[10] Such indicators include (i) the clause-final position of separable particles, negation, and sentential adverbs in main clauses with finite main verbs, and (ii) the pre-verbal position of objects, separable particles, negation, and sentential adverbs in main clauses with modal verbs/perfective *have* and non-finite main verbs.

[11] In Pintzuk 1991, 1993, IPs in Old English are either head-medial or head-final, with obligatory movement of the finite verb to Infl; V2 is analyzed as leftward movement to Infl in Infl-medial clauses. According to this analysis, an increase in the frequency of V2 does not reflect an increase in the use of an optional leftward movement rule, but rather an increase in the use of an Infl-medial grammar.

(18) Swa sceal geong guma gode gewyrcean
So shall young men good-things perform
'Young men shall perform good deeds in this way.'
(Beo 20)

Lightfoot cites Klein (1974) for evidence that Dutch language learners pay attention to Dutch clauses analogous to (18), and Lightfoot (1991: 62, 64) suggests that the order of object and verb in clauses like (18) was accessible to Old English language learners. If the rate of postposition remained constant during the Old English period, with the frequency of clauses like (18) also remaining constant, it seems plausible that these clauses could have been used as evidence for OV structure by children learning Old English. With such a robust indicator of OV structure still in existence at the end of the Old English period, there is no clear support for the hypothesis that the increased frequency of clauses like (16) could have triggered the change from OV to VO.

We can see that although the frequency of VO surface word order does increase during the Old English period, arguments that link this increased frequency and the OV to VO change to an increase in the rate of V2 and/or postposition are not convincing.

5. Synchronic competition between OV and VO grammars

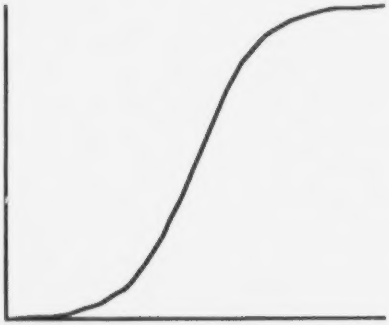
Section 4 presented three types of evidence to contradict the standard account of the change from OV to VO word order at the end of the Old English period. First, clauses unambiguously derived from a VO grammar are used productively during the Old English period, before the change is supposed to have taken place. Second, clauses unambiguously derived from an OV grammar are used productively during the Middle English period, after the change is supposed to have taken place. And third, the increase in VO surface word order during the Old English period and the trigger for change at the end of the period cannot be directly linked to an increase in the rate of either postposition rules or V2.

The evidence points to a different picture of the change from OV to VO. Instead of a uniform grammatical system during the Old English period, with word order variation derived by optional movement rules, there are two competing grammars, one underlyingly OV, the other underlyingly VO. The VO grammar emerges early in the Old English period, and competes with the old OV grammar throughout the Old and Middle English periods, until the old system dies out. Thus the variation in surface word order in both Old and Middle English is at least partially the result of the use of two different grammatical systems, rather than one system with optional rules. And the increase in VO surface word order is at least partially the result of an increase in the use of the new VO grammar, rather than simply an increase in the frequency of use of movement rules.

This analysis replicates the analysis of grammatical competition in languages as diverse as Old French (Kroch 1989), Middle Spanish (Fontana 1993), Old English (Pintzuk 1991, 1993), Middle English (Kroch 1989), Early Yiddish (Santorini 1989, 1993), and Ancient Greek (Taylor 1994). Changes of this type that have been analyzed quantitatively follow an S-shaped curve, as shown in Figure 2: the change starts slowly, accelerates in the middle of the period, and then tapers off to completion.

It should be pointed out that in apparent contradiction to this analysis, many scholars (Gorrell 1895, Kellner 1892, Kohonen 1978, Lightfoot 1991, Mitchell 1985, Stockwell and Minkova 1991) have noticed an abrupt decrease in the frequency of verb-final word order in subordinate clauses at the earliest stages of Middle English, an observation that seems to refute the claim of competing grammars during the Middle English period. But if the change in the underlying order of verbs and their complements is a change of the type shown in Figure 2 above, and if the accelerating middle section of the curve coincides with the end of the Old English period, then a low frequency of OV word order in the Middle English data is only to be expected. Furthermore, it must be emphasized once again that surface word order does not always reflect underlying structure, and that it is necessary to abstract away from verb movement to study verb-complement word order. If we assume that the change from Infl-final to Infl-medial structure was complete early in the Middle English period (Pintzuk

Figure 2
S-shaped curve of syntactic change



1991), then subordinate clauses with finite main verbs will necessarily exhibit VO surface word order, with the verb in clause-medial Infl regardless of the underlying verb-complement word order. As discussed in Section 4.2, in subordinate clauses with auxiliary verbs in Early Middle English documents, Kroch and Taylor (1994) found 26% pre-verbal NPs in West Midlands texts and 29% pre-verbal NPs in Southeastern texts. These frequencies indicate that the order of verbs and their complements in Early Middle English did not significantly differ from the order in Old English, and that the grammars used by speakers during the two stages were much more similar than has previously been suggested.

APPENDIX

ABBREVIATIONS

ÆCHom = Thorpe, Benjamin (ed.) (1844) *The Homilies of the Anglo-Saxon Church*. London: Ælfric Society. Reprinted 1971, New York: Johnson. [volume.page.line]

ÆLS = Skeat, Walter W. (ed.) (1881-1900) *Ælfric's Lives of Saints*. The Early English Text Society, Vols. 76, 82, 94, 114. London: Trübner. [life.line]

ApT = Thorpe, Benjamin (ed.) (1834) *The Anglo-Saxon Version of the Story of Apollonius of Tyre*. London: John and Arthur Arch. [page.line]

Beo = Klaeber, Fr. (ed.) (1950) *Beowulf and the Fight at Finnsburg*. Third Edition. Lexington, Mass.: D. C. Heath. [line]

Chad = Vleeskruyer, Rudolf (ed.) (1953) *The Life of St. Chad: An Old English Homily*. Amsterdam: North-Holland. [page.line]

ChronA = Plummer, Charles (ed.) (1892) *Two of the Saxon Chronicles Parallel*. Oxford: Clarendon Press. [page.line (year)]

ChronB = Thorpe, Benjamin (ed.) (1861) *The Anglo-Saxon Chronicle, According to the Several Original Authorities*. London: Her Majesty's Stationery Office. Reprinted 1964, Kraus Reprint Ltd. [page.line (year)]

ChronE = Plummer, Charles (ed.) (1892) *Two of the Saxon Chronicles Parallel*. Oxford: Clarendon Press. [page.line (year)]

CP = Sweet, Henry (ed.) (1871) *King Alfred's West-Saxon Version of Gregory's Pastoral Care*. The Early English Text Society, Vols. 45, 50. London: Trübner. [page.line]

Or = Bately, Janet M. (ed.) (1980) *The Old English Orosius*. The Early English Text Society, SS, Vol. 6. London: Oxford University Press. [page.line]

REFERENCES

Aitchison, J. (1979) The order of word order change. *Transactions of the Philological Society* 77.43-65.

Allen, Cynthia L. (1975) Old English modals. *Papers in the History and Structure of English*, ed. by Jane Grimshaw, 89-100. (University of Massachusetts Occasional Papers in Linguistics 1.)

Bean, Marian C. (1983) *The Development of Word Order Patterns in Old English*. Totowa, NJ: Barnes and Noble.

Besten, Hans den and Jerold A. Edmondson (1983) The verbal complex in continental West Germanic. *On the Formal Syntax of the Westgermania*, 155-216, ed. by Werner Abraham. (Papers from the 3rd Groningen Grammar Talks, Groningen, January 1981.) Amsterdam: John Benjamins.

Canale, William Michael (1978) *Word Order Change in Old English: Base Reanalysis in Generative Grammar*. Montreal: McGill University dissertation.

Chomsky, Noam (1981) *Lectures on Government and Binding*. (Studies in Generative Grammar 9.) Dordrecht: Foris.

Chomsky, Noam (1986) *Barriers*. (Linguistic Inquiry Monograph 13.) Cambridge, MA: The MIT Press.

Evers, Arnold (1975) *The Transformational Cycle in Dutch and German*. Utrecht: University of Utrecht dissertation. Distributed by the Indiana University Linguistics Club.

Evers, Arnold (1981) Two functional principles for the rule 'Move V'. *Groninger Arbeiten zur germanistischen Linguistik* 19.96-110.

Fontana, Josep M. (1993) *Phrase Structure and the Syntax of Clitics in the History of Spanish*. Philadelphia: University of Pennsylvania dissertation.

Foster, Tony and Wim van der Wurff (1993) The survival of object-verb order in Middle English: some data. Leiden: University of Leiden, ms. *Neophilologus*, to appear.

Foster, Tony and Wim van der Wurff (1994) From syntax to discourse: the function of object-verb order in Late Middle English. Leiden: University of Leiden, ms. Paper presented at the International Conference on Middle English Language, Rydzyna.

Gorrell, J. Hendren (1895) Indirect discourse in Anglo-Saxon. *Publications of the Modern Language Association* 10.342-485.

Haegeman, Liliane (1994) Verb raising as verb projection raising: some empirical problems. *Linguistic Inquiry* 25.509-21.

Haegeman, Liliane and Henk van Riemsdijk (1986) Verb projection raising, scope and the typology of verb movement rules. *Linguistic Inquiry* 17.417-66.

Higgins, F. Roger (1991) The fronting of non-finite verbs in Old English. Paper presented at the annual meeting of the Linguistic Society of America Annual Meeting, Chicago.

Hiltunen, Risto (1983) *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb: The Evidence from Some Old and Early Middle English Texts.* Turku: Turun Yliopisto.

Kellner, Leon (1892) *Historical Outlines of English Syntax.* London: Macmillan.

Kemenade, Ans van (1987) *Syntactic Case and Morphological Case in the History of English.* Dordrecht: Foris.

Kemenade, Ans van (1994) V2 and embedded topicalization in Old and Middle English. Free University Amsterdam/Holland Institute of Generative Linguistics, ms. *Inflection and Syntax in Language Change*, ed. by Ans van Kemenade and Nigel Vincent. Cambridge: Cambridge University Press, to appear.

Klein, R. (1974) Word order: Dutch children and their mothers. *Publikaties van het Instituut voor Algemene Taalwetenschap* 9. Amsterdam.

Kohonen, Viljo (1978) *On the Development of English Word Order in Religious Prose Around 1000 and 1200 A.D.: A Quantitative Study of Word Order in Context.* Meddelanden Fran Stiftelsens for Åbo Akademi Forskningsinstitut, Nr. 38. Publications of the Research Institute of the Åbo Akademi Foundation. Åbo: Åbo Akademi.

Koopman, Willem F. (1990) *Word Order in Old English, with Special Reference to the Verb Phrase.* (Amsterdam Studies in Generative Grammar 1.) Amsterdam: The Faculty of Arts.

Kroch, Anthony S. (1989) Reflexes of grammar in patterns of language change. *Language Variation and Change* 1.199-244.

Kroch, Anthony S. and Beatrice Santorini (1991) The derived constituent structure of the West Germanic verb-raising construction. *Principles* *and Parameters in Comparative Grammar*, ed. by Robert Freidin, 269-338. Cambridge, MA: The MIT Press.

Kroch, Anthony S. and Ann Taylor (1994) Remarks on the XV/VX alternation in Early Middle English. Paper presented at the Third Diachronic Generative Syntax Workshop, Amsterdam.

Lightfoot, David W. (1991) *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: The MIT Press.

Linson, Brian (1993) A pragmatics of word order in Old English prose. Philadelphia: University of Pennsylvania, ms.

Mitchell, Bruce (1985) *Old English Syntax*. Oxford: Clarendon.

Mitchell, Bruce, Christopher Ball, and Angus Cameron (1975) Short titles of Old English texts. *Anglo-Saxon England* 4.207-21.

Mitchell, Bruce, Christopher Ball, and Angus Cameron (1979) Short titles of Old English texts: Addenda and corrigenda. *Anglo-Saxon England* 8.331-3.

Morohovskiy, A.N. (1980) Slovo i predlozenie v istorii angliyskogo yazika. Kiev: Visca skola.

Pintzuk, Susan (1991) *Phrase Structures in Competition: Variation and Change in Old English Word Order*. Philadelphia: University of Pennsylvania dissertation.

Pintzuk, Susan (1993) Verb seconding in Old English: verb movement to Infl. *The Linguistic Review* 10.5-35.

Pintzuk, Susan (1994) Cliticization in Old English. *Second Position Clitics and Related Phenomena*, ed. by Aaron L. Halpern and Arnold Zwicky. Stanford, California: CSLI Press, to appear.

Santorini, Beatrice (1989) *The Generalization of the Verb-Second Constraint in the History of Yiddish*. Philadelphia: University of Pennsylvania dissertation.

Santorini, Beatrice (1993) The rate of phrase structure change in the history of Yiddish. *Language Variation and Change* 5.257-83.

Stockwell, Robert P. (1977) Motivations for exbraciation in Old English. *Mechanisms of Syntactic Change*, ed. by C. Li, 291-314. Austin: University of Texas.

Stockwell, Robert P. and Donka Minkova (1991) Subordination and word order change in the history of English. *Historical English Syntax*, ed. by Dieter Kastovsky, 367-408. (Topics in English Linguistics 2.) New York: Mouton de Gruyter.

Taylor, Ann (1994) The change from SOV to SVO in Ancient Greek. *Language Variation and Change* 6.1-37.

Warner, Anthony R. (1993) *English Auxiliaries: Structure and History.* Cambridge: Cambridge University Press.

SITUATING *QUE*[*]

Bernadette Plunkett

Department of Language and Linguistic Science
University of York

The correct analysis of questions in French is of considerable theoretical interest and much discussion has been devoted to them in the literature on French syntax. One particularly intractable subset of these are 'what' questions. There are various restrictions on these types of questions which, though easy enough to describe are difficult to explain from a theoretical perspective. Of the numerous researchers who have worked on this area (including Obenauer 1976, Goldsmith 1978, Hirschbühler 1978, Koopman 1982, Friedemann 1991, Plunkett 1994) two (Friedemann and Koopman) have explicitly argued that part of the paradigm can be taken to show that certain question phrases are required to undergo Wh Movement into the C projection in the overt syntax of French, even though in other cases such movement can be left until LF. We will see that this is perhaps true, but I will argue that the obligatory movement in such cases can be attributed to independent factors and cannot be taken as proof of a general ban on *in situ* wh-subjects.

In this paper I will redraw the lines around the problematic paradigm and present a new analysis of it. I will then go on to discuss the theoretical implications of the proposed approach.

I begin, in Section 1, by reviewing the relevant facts and summarising the pertinent claims about *que* and *quoi* questions. In

---

[*] This paper owes much to comments on a previous draft by David Adger and Anthony Warner and to lengthy discussions with the former as well as to discussion and judgements from Paul Hirschbühler, Marie-Anne Hintze and Georges Tsoulas. Thanks also to Marie-Laure Masson and Farid Aït Si Selmi for judgements.

York Papers in Linguistics 17 (1996) 265-298
© Bernadette Plunkett

Section 2 I lay out my assumptions about the working of Wh Questions in general and in Section 3 I present the analysis. Section 4, in which the theoretical implications are discussed, concludes the paper.

1. French Questions: Some Restrictions on 'What'

French 'what' questions are special in several respects. Though the final account will link these peculiarities, for the time being I will treat them as separate issues, reviewing each of the restrictions in turn.

1.1 What is 'what'?

Generally speaking, surface Wh Movement is optional in direct questions in French. Wh-words may either move to the front of the sentence or stay *in situ*. A straightforward example of this can be seen in (1).

(1) a. Qui aimes tu?
who love you
'Who do you love?'
b. T(u) aimes qui?

The (b) case here can, but need not be, interpreted as an echo question. The same variability can be seen in the long-distance questions in (2).

(2) a. Qui as tu dit que tu aimes?
who have you said that you love
'Who did you say you loved?'
b. T(u) as dit que t(u) aimes qui?

In fact the two forms may belong to different registers but for most speakers both are possible.[1]

---

[1] Further variability is involved when questions with full noun phrase subjects occur since different types of inversion are available after movement, or indeed no inversion at all. As far as I can tell nothing I have to say about 'what' questions impinges on an adequate account of these different types and I will abstract away from these issues in what follows.

As can be seen, in the case of 'who' questions, the wh-word takes the same form in moved and *in situ* questions. This is not the case in 'what' questions, as (3) shows.

(3) a. Que cherchez vous?
what seek you
What are you looking for?'
b. Vous cherchez quoi?

Not only are there two forms for the word 'what' but they are in complementary distribution, as can be seen in (4).

(4) a. * Vous cherchez que?
You seek what
b. * Quoi cherchez vous?
what seek you

This fact leads to the suggestion, adopted by most researchers in the area, that they are variants of the same morpheme (but see Obenauer 1976, 1977 for a different view). On this view the two forms of the word for 'what' may be seen as a weak unstressed form *que* and a tonic form *quoi*. This view is supported by the fact that the variants are similar to those found in other weak-strong pronominal pairs such as *te ~ toi, me ~ moi, se ~ soi*. It is further supported by the fact that, just as with those pairs, only the strong form appears inside PPs:

(5) a. Vous pensez à quoi?
you think to what
'What are you thinking about?'
b. A quoi pensez vous?
c. * Vous pensez à que?
d. * A que pensez vous?

In addition, for most speakers *que* cannot be co-ordinated with another wh-word. Thus (6a) and (6b) are parallel to (6d) where the co-ordination of weak subject pronouns is ruled out, while (6c) is perfect.

(6) a. ?? Qui ou que voulez-vous photographier?
who or what want you to photograph
'Who or what do you want to photograph?'
b. * Que ou qui voulez-vous photographier?
c. Qui ou quoi voulez-vous photographier?
d. * Tu et il voulez photographier quelqu'un.
'You and he want to photograph someone.'

The treatment of *que* as a weak form of *quoi* then is well-supported, but as we will see below the precise characterisation of 'weak' pronouns is somewhat problematic.

The alternative view of the alternation in (3) is the one put forward by Obenauer in which *que* in fronted questions is treated as the finite complementiser *que* while *quoi* is treated as a genuine wh-word. This treatment parallels that of Kayne (1976) and others for the *que* which appears in relative clauses. However, while accepting Kayne's analysis for relative *que*, both Goldsmith (1978) and Hirschbühler (1978) review and argue in detail against Obenauer's view of interrogative *que*. Their arguments are convincing; for example, as Goldsmith (1978, 1981) points out, simple inversion of a verb and a pronominal subject is blocked by the presence of an overt complementiser, not only in embedded clauses in French but in matrix clauses too in the cases where a complementiser may appear in them.

(7) a. Peut-être qu'il est parti.
perhaps that he is left
'Perhaps he has left.'
b.* Peut-être qu'est-il parti.
perhaps that is he left
c. Peut-être est-il parti.
perhaps is-he left

Since this type of inversion does take place in interrogatives, as we have seen in (1-3), the *que* there cannot be a complementiser unless just in this case the verb is allowed to raise to C and adjoin to the right of the overt complementiser. If this were to happen then clearly the *que* complementiser in (3) and the *que* complementiser in (7) would have to be differentiated from one another. In fact, to the extent that *que* must always appear immediately before the inflected verb and any clitics it may have attached to it, as claimed by Obenauer (1977), all *que* questions containing pronominal subjects will involve simple inversion.[2] Since inversion is typically taken to indicate that the verb is in C, which is borne out by the contrast in (7), it is fairly safe to assume that when *que* appears it is always outside IP.

It would seem then that the two views on the status of interrogative *que* are incompatible. However, within current syntactic analyses couched in the Principles and Parameters framework they can be seen to have something in common. Complementisers and pronouns are both treated as functional heads which may have syntactic complements but do not assign theta roles and hence cannot take arguments. Since this is the case, some aspects of the behaviour of *que* may be attributed to its status as a functional head and are thus compatible with its treatment as a pronoun in the current framework in a way which was not possible in earlier approaches.

1.2 Subject questions

Further and yet more problematic constraints on 'what' arise in that in simple direct questions if it functions as the subject it appears neither to be possible to extract it, nor (if we take *quoi* to be the form used when it has not been moved) to be able to stay *in situ*.

(8) * Que flotte dans l'eau?
what floats in the water
'What floats in water?' or 'What is floating in the water?'

(9) * Quoi flotte dans l'eau?
what floats in the water

---

[2] Apparent exceptions to this generalisation, like (i), where complex inversion has taken place, are rejected by Obenauer (1977) as marginal but uniformally accepted by my informants.

(i) Que cela veut-il dire?
what that wants it to say
'What does that mean?'

This is not true for other wh-phrases as (10) shows.

(10) Qui flotte dans l'eau?
who floats in the water
'Who is floating/floats in the water?'

The restriction on extraction is not seen in more complex questions like (11), which I take (*pace* Obenauer 1976) to be cases of long-distance extraction given the standard *que* ~ *qui* alternation which shows up after extraction of an embedded subject.[3]

(11) Qu'est ce qui flotte dans l'eau?
what is this that floats in water
'What (is it that) floats/is floating in (the) water?'

These cases completely parallel other cases of long-distance subject-*que* extraction such as (12).

(12) Que crains-tu qui soit advenu?
what fear-you that is taken place
'What do you fear has happened?'

Whether the restriction on *quoi* in [Spec,IP] extends to embedded contexts is harder to determine. The impossibility of cases like (13) suggests that it does.

(13) * Tu pensais que quoi trainait dans le couloir?
you thought that what lay around in the corridor
'What did you think was lying around in the corridor?'

However, an example given to me by Paul Hirschbühler shows that where movement is independently blocked, 'what' may perhaps stay in subject position.

---

[3] In contexts where *that*-t effects would show up in English a *que* complementiser becomes *qui*; the effect is dubbed 'masquerade' by Kayne (1976) and is considered by Rizzi (1989) to be a case of agreement in Comp, with the C showing the presence of a wh-trace in its specifier.

(14) Qui a dit que quoi trainait où?
who has said that what lay around where
'Who said what was lying around where?'

This suggests that the ban on *quoi* in subject position is not merely due to its incompatibility with nominative Case, as Goldsmith (1981) claims.[4] In Plunkett (1994) this explanation for the absence of *que/quoi* subject questions was adopted and it was argued that stressed subject pronouns such as the ones in the echo questions in (15) noted by Koopman (1982) be taken to be non-nominative forms.[5]

(15) a. QUOI a été décidé?
what has been decided
'WHAT was decided?'
b. QUOI flotte dans l'eau?
what floats in the water
'WHAT floats in water?'

Another set of examples which might be problematic for Goldsmith's view are those like (16) where, under most views, the expletive subject would transmit nominative Case to *quoi* in post-verbal position.

(16) Il est arrivé quoi?
it is happened what
'What happened?'

These arise both with unaccusative type verbs such as those which occur in English *There*-Insertion constructions and, in French in passives, as in

---

[4] Though that approach has the advantage of being able to explain why many speakers only marginally accept *quoi* in subject positions in echo questions and others reject it altogether.

[5] It was felt that the contrast in (6) supported that view.

(17) Il a été décidé quoi pour demain?
it has been decided what for tomorrow
'What has been decided for tomorrow?'

These types of construction provide additional information about the constraints on the extraction of 'what' since, when [Spec,IP] is filled with an expletive, the post-verbal nominative *que* can be extracted as (18) and (19) show.

(18) Qu'est-il arrivé?
what is it happened
'What happened?'

(19) Qu'a-t-il été décidé pour demain?
what has it been decided for tomorrow
'What has been decided for tomorrow?'

This possibility might lead us to wonder whether the cases of apparent long-distance subject *que* movement in (12) were not in fact instances of extraction from a post-verbal position, since native speakers often have difficulty in deciding which of the examples in (20) is the appropriate way of writing the corresponding spoken question.

(20) a. Que dis-tu qui est advenu?
what say you that is happened
'What do you say happened?'
b. Que dis-tu qu'il est advenu?
what say you that it is happened

However, there are clear cases where no expletive subject is possible, as in (21) and long distance subject extraction is indeed still licit.

(21) Que prétendais-tu qui motivait cette analyse?
what claimed you that motivated that analysis
'What did you claim motivated that analysis?'

One might wonder whether any further information could be gleaned from looking at indirect subject questions. Unfortunately, this is not possible. 'What' questions in this context are in fact anomalous, but in this case, as the paradigm in (22) shows, there is no difference between subject questions and object ones; when the embedded clause is tensed, neither permits a simple question introduced by *que*. Instead, these indirect 'what' questions are always introduced by the pronoun *ce* ('it'), resulting in a free-relative type structure.

(22) a. * Je me demande que/quoi tu aimes.
I myself ask what you like
b. Je me demande ce que tu aimes.
I myself ask it that you like
'I wonder what you like'
c. * Je me demande qui/quoi lui fait peur.
I myself ask what him makes frightened
d. Je me demande ce qui lui fait peur.
I myself ask it that him makes frightened
'I wonder what makes him frightened.'

This restriction is specific to indirect 'what' questions, since the instances of (23) are unexceptional.

(23) a. Je me demande qui tu aimes.
I myself ask who you like
'I wonder who you like.'
b. Je me demande qui lui fait peur.
I myself ask who him makes frightened
'I wonder who makes him frightened.'

The restriction could be linked to the dependence of *que* on an adjacent verb but it can have nothing to do with the status of subject questions. In fact, in the questions in (22b) and (d) *que* is clearly the relative complementiser as Kayne (1976) argued was the case in all relatives, since where the subject has been extracted we find the *qui* alternant though the head of the relative is inanimate.

Where the wh-clause is non-finite the facts are different again but since in these cases there can never be an overt subject they cannot be relevant with regards to the restriction on subject questions.[6] Since it is that restriction which I will now concentrate on, in what follows I will abstract away from indirect questions.

1.3 Review

We have seen that *que/quoi* questions are special in several ways. First, 'what' has two forms in French, one appearing to be a weak or clitic pronoun which undergoes movement and the second a strong pronoun which appears when the *in situ* strategy for Wh Questions is used. Second, in matrix direct questions *que* cannot appear bearing the grammatical function of subject, suggesting in by now traditional terms that 'extraction' of 'what' subjects is impossible in French. However, coincidentally *quoi* may not appear as an *in situ* matrix subject either and it is unclear how closely these facts should be related to the availability of two forms for the 'what' pronoun.

In the next section I will be discussing one approach to Wh Movement with a view to seeing whether it can shed any light on these peculiarities.

2. Wh Movement

Rizzi (1991), reformulating the approach taken in May (1985), proposed that Wh Movement could be accounted for by the Wh Criterion as given in (24).

---

[6] In infinitivals (as discussed in Hirschbühler 1978) we find the only case where *que* and *quoi* are not in complete complementary distribution. An embedded case is illustrated in (i).

(i)a. Je ne sais quoi faire
I not know what to do
b. Je ne sais que faire
I not know what to do
'I don't know what to do'

Hirschbühler argues that subtle semantic factors distinguish these two.

(24) **Wh Criterion**

a. A Wh-operator must be in a Spec-head configuration with an X^0_{+WH}

b. An X^0_{+WH} must be in a Spec-head configuration with a Wh-operator

(Rizzi 1991: 2)

In Plunkett (1993) a similar, if somewhat less strict, approach is taken with regard to questions where the principle in (25) is essentially comparable to clause (b) of the Wh Criterion.[7]

(25) **Interrogative Movement Principle** (IMP)

The specifier of a head which bears question features must bear matching features.

(Plunkett 1993: 262)

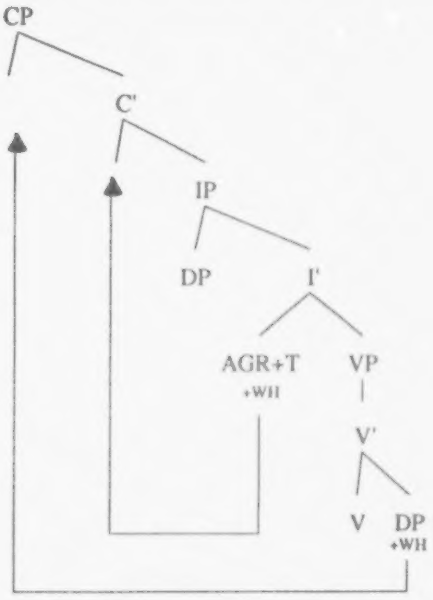
Although the two approaches diverge in detail, they converge in the proposal that wh-features are marked on C in selected embedded wh-clauses but on the head which is normally immediately below C in root clauses; we also agree that the principle applies at S-structure in English. Rizzi assumes that in root clauses wh-features are associated with the head containing tense features whereas I located them in Agr; these details seem to be irrelevant to the analysis of the French data and for the sake of simplicity I will illustrate with a unified Infl assuming this to contain both Tns and Agr features.

The complementarity between inversion in root and embedded clauses in English questions has led to the now standard analysis of [Spec,CP] as the landing site for Wh Movement. Although both approaches situate wh-features lower than C in root clauses, the claim that [Spec,CP] is the usual landing site for Wh Movement is not disputed. Both approaches employ the same mechanism to explain why a wh-phrase usually ends up in [Spec,CP] in English; the subject occupies [Spec,IP] so that the principle in (25) usually cannot be satisfied by S-structure unless I moves into C, whose specifier is empty; the wh-phrase can then move into the specifier position, permitting spec-head agreement in the C projection with respect to wh-features. A typical pre-Wh Movement structure would be the one shown in (26), where arrows show the subsequent movement.

---

[7] Clause (a) in (24) was originally intended to deal with non-inverting structures such as relative clauses and will not be of relevance until Section 3. In the meantime I will refer only to the IMP in (25) with the understanding that in nearly all cases, (24b) and (25) have the same coverage.

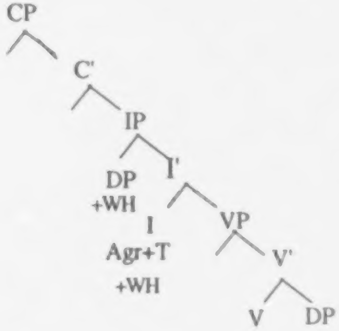
(26)



The Infl node and the subject NP do not agree in wh-features; if, however, both the object NP and the head marked +wh, move into the C projection then the IMP will be satisfied. The same type of situation will arise when an adjunct phrase or an argument in a lower clause is marked +wh.

There is one type of construction, however, where the approaches differ more substantially; this is the configuration in which the root subject is marked +wh, as in (27).

(27)



In a configuration such as the one in (27), IMP is immediately satisfied. I assume the now familiar Lexical Clause Hypothesis with subjects in French and English raising to [Spec,IP] to get Case; the subject and Infl agree in wh-features here and there is no obvious motivation for further movement of either the wh-phrase or the wh-marked head. Since this is so, considerations of economy would lead us to expect that no further movement of the wh-phrase will be required either in the syntax or at LF; indeed, I will argue not only that further movement is unnecessary but that once IMP has been satisfied, it is impossible. In so far as this approach requires the minimum number of steps it is the ***Minimal Approach*** to Wh Movement and will be referred to as such in what follows. Rizzi (1991) acknowledges that to say that no further movement takes place in such cases is the most straightforward account of root subject questions in English. The analysis correctly predicts that we will see no evidence of Subject Auxiliary Inversion in such questions, this being a movement which is triggered to allow satisfaction of the IMP. While the absence of inversion in such questions is an effect which people have previously struggled to explain, it is a natural consequence of the Minimal Approach.

However, Rizzi (1991) does not adopt the Minimal Approach. One of his reasons is that part of the data from French questions, discussed in the previous section, can be taken to indicate that subject wh-phrases must vacate [Spec;IP]. As mentioned in the introduction, this was the conclusion reached on different grounds by both Koopman (1982) and Friedemann (1991). In the following section I will discuss the analysis of French questions with respect to the type of approach outlined, first in general and then with respect to the specific restrictions on 'what' questions. As far as subject questions are concerned I will focus on how the Minimal Approach can cope with the French data.

3. The Minimal Approach to French Questions

An adequate approach to Wh Movement must be able to account for when any wh-phrase must, may or may not move. In addition, it should correctly predict in which cases of Wh Movement a concomitant inversion must or may take place. In particular, leaving aside factors specific to subject questions for the moment, with respect to French it must explain:

(i) why (overt) Wh Movement is optional in matrix questions and obligatory in embedded questions;
(ii) why inversion is possible but not obligatory with most matrix (moved) questions but impossible in embedded questions;[8]
(iii) why, in obligatory contexts, only one wh-phrase has to move;
(iv) why inversion never happens when a wh-phrase stays *in situ*;
(v) why partial Wh Movement is not possible (eg. movement to an intermediate [Spec,CP]).

In addition, with respect to 'what' questions, our theory must explain:
(vi) why inversion is obligatory in matrix *que* questions.

Rizzi (1991) deals with the first five of these. I will begin my analysis by looking in detail at these factors and propose some modifications to his treatment. Next, I will turn to the treatment of 'what' questions specifically and finally, I will discuss subject questions in general and argue that we should ensure that the approach is 'minimal'.

---

[8] Stylistic Inversion is sometimes found in embedded contexts and is thus an exception to this generalisation. A full investigation of the differences in different types of inversion is beyond the scope of this paper.

3.1 Optional inversion and optional movement

As we saw above, the IMP in (25) has the same effect as clause (b) of the Wh Criterion (24) which was designed to deal with inversion constructions. Let us first examine how the inversion data is explained and then proceed to look briefly at non-inversion in French questions and whether clause (a) of (24) or an equivalent is also necessary.

If the head of every question clause bears wh-features and, if (24)/(25) applies at S-structure in French (as Rizzi (1991) claims), then Wh Movement should be obligatory, as it is in English. This is a correct prediction for indirect questions in French, where the matrix verb selects a CP whose head is marked +wh, but since matrix Wh Movement is optional Rizzi proposes that while matrix I may bear wh-features, such features are not *necessarily* generated. He points to the optionality of the question marker *ka* in Japanese matrix questions in support of this claim.[9] This proposal that wh-features are generated freely, which largely accounts for factor (i), seems reasonable and I will assume in what follows that in a direct question where no wh-phrase moves, the head of the matrix clause is -wh. The question now arises whether obligatory Wh Movement indicates that all question clauses must obligatorily have +wh heads in English. It would seem rather *ad hoc* to assume that wh-features are freely generated in French but obligatorily generated in certain contexts in English. However, another of his proposals allows Rizzi to circumvent this problem. In positing two clauses of the Wh Criterion Rizzi is in effect postulating that spec-head matching in wh-features is required independently by both wh-heads and wh-phrases. This entails that when the head of an unselected clause is -wh but the sentence contains a wh-phrase, Wh Movement will still be required at some level, as has usually been assumed. Rizzi argues that when this situation arises in French, the wh-phrase may move overtly to [Spec,CP] then, by a process of 'dynamic agreement' the empty C position will come to agree with the wh-phrase and (24a) will be satisfied. In this case, since no wh-feature has been forced to move from I to C, no inversion will take place and Rizzi thus explains factor (ii) which accounts for the possibility of uninverted questions like (28) in French.

---

[9] It is obligatory in embedded questions.

(28) Comment tu l'as su?
how you it have known
'How did you know that?'

Rizzi (1991) argues that English lacks Dynamic Agreement. Since, on his view, a question with no wh-head would only be able to satisfy (24) if Dynamic Agreement were available, the postulation that it exists in French but not English will account both for the fact that all questions involve both overt movement and inversion in English.

Now, Rizzi (1991) assumes that both clauses of the Wh Criterion (24) must apply at the same level in a given language, thus incidentally explaining factor (iv), i.e. why clause (b) cannot be satisfied simply by the operation of inversion, with subsequent movement of the wh-phrase left until LF. However, if the presence of a wh-phrase is itself sufficient to cause movement, as clause (a) of (24) suggests, then the possibility that no +wh head will be generated in a given matrix context ought not to be sufficient to predict the possibility of *in situ* questions in French. In Rizzi (1991) the explanation for the fact that some wh-phrases can remain *in situ* until LF is maintained by the additional assumption that these do not have the status of 'operators' until that level and, as a result, clause (a) does not apply to them until then. [10]

Overall then, Rizzi's (1991) approach manages to account for all the factors in (i) to (iv) above but under current economy considerations the approach faces a problem. If wh-phrases are not deemed to be operators until LF, an assumption required for English where factor (iii) also holds, why should they (be able to) move in the syntax in cases like (28) in French? Economy predicts that even if Dynamic Agreement were available it should only ever be invoked at LF.

---

[10] An alternative explanation of this option which Rizzi considers and rejects is that clause (a) of (24) not apply until LF in French. Although indirect questions (and relative clauses) involve no inversion, clause (b) is sufficient to ensure obligatory movement in them. Late application of clause (a) would have the desired effect then of correctly predicting not only the possibility of *in situ* questions but also giving an account of factor (iii), why in multiple wh-questions in French as in English only one wh-phrase may move in the syntax.

Under Minimalism (Chomsky 1993, 1995), pure optionality of movement is ruled out. Movement of an element in the syntax is licit only if a failure of such movement would result in a derivation which could not converge. With a view to explaining the French data within the current approach while retaining as much of the explanatory power of Rizzi's approach as possible I would like now to propose some revisions.

Let us assume as before that wh-features are generated freely in unselected environments. If none of the clausal heads have been generated with wh-features but a sentence contains a wh-phrase, then that phrase will be **required** to stay *in situ*. However, semantic requirements will mean that unless the scope of the wh-phrase can be determined in some other way the sentence will be uninterpretable. Leaving aside details, let us assume that languages which allow *in situ* wh-phrases have access to such a mechanism while languages like English do not. On such a view, visible movement entails the presence of wh-features on some clausal head while lack of movement entails the absence of such features. If this is correct then an alternative explanation for uninverted structures like (28) must be sought. Consider for a moment what form such structures take in the varieties of French in which the Doubly Filled Comp Filter (DFCF) (Chomsky and Lasnik 1977) is not in effect.

(29) Comment que tu l'as su?
how that you it have known
'How did you know that?'

One might claim that the C here is -wh and invoke something like Dynamic Agreement in such structures, but given that it will be necessary to assume that in these dialects C can be freely generated in root contexts, it is much more straightforward to assume that when C is the head of the clause, that is the head that any wh-features will appear on. If Dynamic Agreement is not involved in (29), some head bears wh-features or the Wh Criterion could not be satisfied; it must be C or an *ad hoc* mechanism will be required to explain the grammaticality. Now suppose that with respect to Wh Questions, dialects such as Metropolitan Standard French (MSF) and Québécois differ only in their application of the DFCF. If wh-features can be generated on C in root clauses in French,[11] the operation of the DFCF in some dialects will explain the absence of an overt complementiser in cases like (28) but the presence of a non-overt +wh complementiser there will obviate the need for inversion and its absence will thus be explained. It would be superfluous to assume that the dialects differ further by invoking Dynamic Agreement for cases such as (28). Since in MSF movement with inversion is also possible we need only claim that the projection of C is optional in French root clauses. This claim is independently supported by the following well known contrast seen in example (7) in which either inversion or an overt complementiser is possible after certain sentential adverbs in MSF, but not both.

(30) Peut-être est-il parti.
perhaps is he left
'Perhaps he left.'

(31) Peut-être qu'il est parti.
perhaps that he is left

(32) * Peut-être qu'est-il parti.
perhaps that he is left

If this approach is correct and Dynamic Agreement can be dispensed with in the explanation of structures like (28) then what accounts for the absence of uninverted questions in English? The simplest account must be correct here: complementisers cannot be generated in matrix contexts in English.

---

[11] We must ensure that the DCFC operates only in wh-contexts in which the C projection is filled with a complementiser and not when it is filled with a verb, i.e. when the C position is filled at D-structure.

Having dispensed with the need for Dynamic Agreement in non-inversion structures the question now arises of whether it is needed at all. Under standard GB assumptions, in multiple questions where one wh-phrase has moved in the syntax, movement of any remaining wh-phrases involves absorption (Higginbotham and May 1981), clause (a) of (24) or its equivalent presumably being responsible for the movement. Suppose however, that LF movement of an *in situ* wh-phrase is required merely because of the need for scope assignment rather than because of an independent spec-head requirement on wh-phrases as such. Since the presence of the word 'operator' in (24a) is crucial to an adequate description of the data it is unclear that this clause can be in operation for anything other than semantic reasons. If this is the only motivation for the postulation of clause (a) and its effect can be guaranteed by independent requirements, then it should be dispensed with, leaving a single-pronged Wh Criterion. Such a version of the criterion would be much more in keeping with Chomsky's recent proposals concerning the operation of Checking Theory (Chomsky 1995). Suppose then that there is no clause (a) to the Wh Criterion and that *in situ* wh-phrases may be assigned scope by some means other than movement and absorption at LF. If this approach is correct then there will be no need to invoke Dynamic Agreement at LF and it can thus be dispensed with completely.[12]

Before proceeding, let us look briefly at whether the proposed revisions to Rizzi's approach explain factor (v), the lack of partial Wh Movement in French and English and continue to allow us to explain factor (iv), why we never find inversion without concomitant Wh Movement.

Under the monoclausal approach to the Wh Criterion there is a one-to-one correspondence between the presence of a wh clausal head and the application of overt Wh Movement. Once the head of an IP or CP has +wh-features, the revision of (24b)/(25) in (33) will kick in.

---

[12] The question of what precisely happens to unmoved wh-phrases at LF is left open here. In Baker (1970) and indeed in much recent work (Aoun and Li 1993, Kiss 1993, Stroik 1995, Williams 1986) LF movement is not invoked to explain the assignment of scope to *in situ* wh-elements.

(33) **Wh Criterion** (revised)
Heads marked +wh bear a strong (alternatively weak) (Categorial) X feature[13]

Since strong features must be eliminated by Spell-out (S-structure), it follows that partial movement should never be licit in a language in which the categorial feature on a wh-head is strong.[14] Under a checking view of the Wh Criterion it follows too that inversion could never take place without concomitant Wh Movement. Factors (iv) and (v) then fall out quite neatly within this framework.

Before proceeding to the next section in which we consider factor (vi) let us briefly summarise the assumptions entailed in the revised approach to Wh Movement taken here.

In unselected contexts wh-features are freely generated on a clausal head. Some languages limit the choice of clausal head in root contexts (English) while others allow a choice between the projection of an inflectional head only or a complementiser (French). Where a choice is available, wh-features may freely appear on the topmost head; where this is a head such as C which unlike I does not independently require its spec to be filled, uninverted questions will be possible. These may be of two types: those like spoken MSF in which the DFCF operates and those like Québécois in which it does not. (Visible) Wh Movement is triggered solely by the presence of a strong categorial feature on any wh-marked head, which in French may be either I or C. There is an isomorphic relation between the presence of a clausal head marked +wh and Wh Movement. In some languages assignment of scope to a wh-phrase at LF is limited to contexts in which a wh-phrase has already moved in the syntax, so that in these languages all derivations of questions in which no clausal heads are marked +wh will crash; English is such a language while French is not. Note that it is with respect to the presence or absence of this mechanism that English and French are postulated to differ rather than with respect to Dynamic Agreement.

---

[13] Where an X feature is similar to a D-feature as in Chomsky (1995) but where clearly the particular category of the element is unimportant.

[14] How languages such as those described in McDaniel (1989) should be treated is as yet unclear to me.

The proposed revisions are necessary to a complete explanation for the behaviour of 'what' questions in French to which we now return.

3.2 *Que* questions

We begin our re-examination of *que* questions by looking at the reasons for the obligatory inversion which it induces, we then move on to look at the clitic-like nature of *que*.

3.2.1 Obligatory inversion in *que* questions

Let us look again at factor (vi), why 'what' questions always induce inversion in French. Rizzi (1991) did not attempt to deal with this matter, but within both his framework and our revisions of it inversion occurs only where an inflectional head bears +wh-features; we may thus see this restriction as one which rules out derivations in which wh-features are generated on C.[15] As can be seen from the examples in the previous sub-section, when matrix C occurs overtly in French it has the same form as the complementiser which introduces finite embedded clauses, *que* (or *qui* when subject extraction has taken place). We may say then that when the complementiser *que* bears wh-features, movement of the weak form *que* causes the derivation to crash.[16] One might posit a fairly superficial reason why *que* questions are licit only when I bears wh-features such as a filter blocking *que* in the spec of a *que* Comp. The restriction is in fact more likely to have something to do with the clitic-like properties of the question-word *que*, however. One reason is that such a filter would be likely to have a phonological basis and yet in this case we would have to say that it operates even in MSF where the DFCF means that the second of two adjacent *que*s is not even pronounced. The second reason is that a similar situation in which *qui* occupies both the head and spec of CP results in no ungrammaticality in the dialects in which DFCF does not operate.[17]

---

[15] Absence of wh-features is still licit since *quoi* may remain *in situ*.

[16] Note that even in Québécois where there is a clear preference for situating wh-features on C rather than I, when *que* is used inversion must occur.

[17] The complementiser *qui* is not only possible here but according to Lefebvre (1982) it is obligatory for reasons having to do with the ECP.

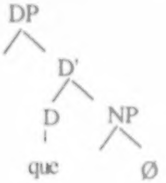
(34) Qui qui est venu?
who that is come
'Who came?'

Since *qui* does not have clitic-like properties this contrast is to be expected if we attribute the restriction to the clitic nature of *que*.[18] Let us explore further the clitic-like nature of the wh-word *que*.

3.2.2 *Que* as a defective clitic

We saw in Section 1 that there are sound morphological and syntactic reasons for regarding *que* as a weak form of the pronoun *quoi*. We may take pronouns to be determiners which head a projection containing a zero nominal head as in (29), and if 'what' in French is a pronoun then we will expect it to sometimes behave as a full phrasal projection (i.e. DP) and sometimes as a head (D).

(35)



---

[18] Further support can be found from the fact that in some dialects of Canadian French the non-clitic form *quoi* may appear in a fronted position, as in (i).

(i) Quoi c'est que Jean fait?
what it is that Jean does
'What is Jean doing?'

Indeed, a few speakers seem to even accept cases like (ii) though Lefebvre (1982) claims that the majority of her informants rejected such cases.

(ii)(*)Quoi tu fais?
what you do
'What are you doing?'

However, I have no explanation for why it is possible to move the strong form alone in these dialects but not in the MSF example in (iii).

(iii)*Quoi fait Jean?
what does Jean
'What is Jean doing?'

A most natural corollary of this view would be to treat *que* as the form which is used when head movement has taken place and *quoi* as the full DP form. This is the view espoused in Plunkett (1994) and it could clearly account not only for the dependent status of *que* but also for the fact that it cliticises only to verbs rather than whatever it happens to be adjacent to. However, adopting this view is not straightforward; weak object pronouns in French are standardly treated as syntactic clitics and since Kayne (1975) clitic placement has been largely regarded as involving movement of a head.[19]

Hirschbühler (1978), advocating a pronominal treatment of interrogative *que*, already argued that it was a clitic, thus accounting for its appearance adjacent to a verb.[20] However, the rules which he invoked to account for its status as a 'dependant' were phonological. While the distribution of *que*, as described by Hirschbühler, clearly shows that it is a phonological clitic on the verb, its status as a syntactic clitic and hence as a head which has undergone head movement is less certain. In particular, as already noted by Friedemann (1991), the fact that *que* can occur in long-distance questions where it has been extracted out of a tensed clause casts strong doubt on the possibility that it reaches the head of the matrix clause by Head Movement, especially since such Long Head Movement is otherwise unknown in French.

---

[19] In more recent approaches movement of a clitic is claimed to take place in two steps, the first, movement of a maximal projection to the specifier of an agreement phrase to get case and the second a further movement of the head to the clitic position. This is the approach I believe to be correct; however, some researchers (eg. Sportiche 1994), base generate clitics in a fronted position.

[20] Aside from the cases mentioned in an earlier footnote, the only exceptions to the requirement that *que* be left-adjacent to a verb involve instances of *que diable* ('what the devil') which is not as restricted in its occurrence as simple cases of *que*. Like *que* this cannot occur next to a subject pronoun.

(i) * Que diable tu cherches?
what devil you look for
'What the hell are you looking for?'

Hirschbühler (1978) points out that all wh-*diable* phrases induce simple inversion.

Suppose we treat *que* as a phonological clitic but not a syntactic one. In this case we could assume that Wh Movement of 'what' in French involves movement of the whole DP until the target position has been reached. At that point the head could pro-cliticise to the adjacent verb or other clitic, where inversion has taken place. This would explain why *que* consistently appears outside all other clitics, including *ne*. It would also enable us to account for the fact that unlike other clitics *que* need not attach to the verb of its own clause, as in (12) and (21) repeated in (36).

(36) a. Que crains-tu qui soit advenu?
what fear you that is taken place
'What do you fear has happened?'

b. Que prétendais-tu qui motivait cette analyse?
what claimed you that motivated that analysis
'What did you claim motivated that analysis?'

(37) Que ne faudrait-il jamais faire t?
what NE ought-it never to do
'What ought one never to do?'

This solution does not require that we invoke Long Head Movement. However, the problem remains of how to account for why it always cliticises to a verb group and never anything else and in particular, why it cannot cliticise to a complementiser. In fact, under the view presented here it is this last case which it is essential to rule out since uninverted questions are posited to contain a non-overt complementiser adjacent to the wh-phrase. Clearly, it will be necessary to assume that phonological clitics like *que* may cliticise only to heads which are structurally adjacent and that these must have phonological content. I would like to propose that what is at stake in the **que que* sequence is that the complementiser does not itself have enough phonological weight to act as a host for a phonological clitic while a verb, plus or minus verbal clitics does.[21]

Assuming that *que* questions in which the matrix C bears wh-features can be ruled out in this way, let us turn now to the remaining problematic cases in which *que* functions as a subject.

3.2.3 *Que* and subject questions

Let us return finally to the restriction on matrix clauses with 'what' subjects in French. As we saw earlier, these appear to be both banned from staying *in situ*, in the [Spec,IP], taking the form *quoi* and from moving to [Spec,CP] and taking the form *que*. Let us now see how this can be explained. To begin, let us review some of the problematic cases:

(38) a. * Que/quoi a été décidé?
what has been decided
'What was decided?'
b. * Que/quoi flotte dans l'eau?
what floats in the water
'What floats in water?'

Simple matrix questions are ungrammatical when the subject is a form of 'what', both when the subject is left *in situ* and when it is moved. However, the echo version of the *in situ* question is acceptable, as we saw in (15), repeated here as (39).

(39) a. QUOI a été décidé?
what has been decided
'WHAT was decided?'
b. QUOI flotte dans l'eau?
what floats in the water
'WHAT floats in water?'

---

[21] Although complement clitics are themselves phonologically light they form a phonological phrase with the following verb. However, phonological weight might also be relevant in accounting for the fact that many speakers find *que* questions where the first clitic is *ne* to be odd.

The impossibility of (38) cannot be attributed to any thematic restriction on *que/quoi* as the thematic relations are the same in (38) as in (39) and presumably they are the same again in the relevant part of (17) repeated here as (40).

(40) Il a été décidé quoi pour demain?
it has been decided what for tomorrow
'What has been decided for tomorrow?'

Note that here the wh-phrase does not occupy the subject position, which is filled instead by an expletive. In addition, we cannot maintain that *que/quoi* simply cannot be a subject because in elliptical questions with no verb *quoi* can clearly refer to the subject as (41) (from Léard 1982) shows.

(41) a. Quelque chose me chagrine.
something me upsets
'Something is upsetting me.'
b. Quoi donc?
what then
'What?

In addition, we have just seen cases in (36) where *que* has been extracted from the subject position in a lower clause. The acceptable periphrastic forms such as the one in (11) repeated here as (42) were taken to fall into this category too.

(42) Qu'est ce qui t flotte dans l'eau?
what is this that floats in the water
'What (is it that) floats/is floating in the water?'

Echo interpretations aside, the contrasts seem generally to show that *que/quoi* may occupy [Spec,IP] but not at S-structure and that *que* may occupy [Spec,CP] but not if it has been extracted from the subject position of the same clause. Let us dispense with the latter case first. Given that 'what' cannot be completely barred from the specifier position of a tensed CP we need to explain why it is blocked from moving the short distance shown in (43).

(43) $* [_{CP}\text{que}_i [_{C} V_j [_{IP} t_i \; t_j \ldots]]]$

This configuration could perhaps be ruled out as an ECP violation which cannot be salvaged by Masquerade, as it can in the embedded clause in the relevant cases, since only IP has been projected. However it is not clear why an inverted verb would not be able to govern the trace position as Rizzi assumes happens with the extraction of a 'who' subject in (44).

(44) Qui vient?
who comes
'Who is coming?

I would like to maintain, though, that the verb has nothing to salvage in (44) since *qui* is in [Spec,IP] and not [Spec,CP]. This is exactly what the Minimal Approach to Wh Movement (as in Plunkett 1993) would predict. Put into the framework presented here, economy considerations will block an I marked +wh from moving to C in this situation since the wh-phrase in its specifier satisfies the revised Wh Criterion in (33) and further movement, being completely unmotivated, is blocked.[22] If movement is blocked in (44) then the same applies in (38), economy thus rules out the representation in (43). It is interesting to compare (18) and (19) repeated here as (45) and (46) in this regard.

(45) Qu'est-il arrivé?
what is it happened
'What happened?'

---

[22] Under Minimalism, movement is permitted only to satisfy morphological requirements and never in order to salvage ungrammaticality.

(46) Qu'a-t-il été décidé pour demain?
what has it been decided for tomorrow
'What has been decided for tomorrow?'

In cases such as these *que* is in fact an underlying object and at S-structure [Spec,IP] is filled by an expletive. In this situation of course economy will not block further movement because the only way to satisfy the Wh Criterion (33) will be for I to move to C and for the wh-phrase to move into [Spec,CP].

Let us concentrate then on explaining the remaining problem, the ban on (non-echo) *quoi*. when *in situ*. I would like to attribute this to the status of *que/quoi* as a non-specific indefinite.[23] Not all of the ungrammatical examples with *quoi* subjects have grammatical equivalents with expletive subjects but it is significant that in the examples usually cited *quoi* is the surface subject of a predicate with a single argument, plausibly an unaccusative,[24] or of a passive predicate. In fact, when we look at a different type of predicate speakers will sometimes, at least marginally, accept *que* subjects. The following have been found acceptable by more than one speaker.

(47) ? Que démontrait le redressement de l'économie?[25]
what demonstrated the re-establishment of the economy
'What demonstrated the recovery of the economy?'

---

[23] My thanks go to David Adger for first suggesting to me that the contrast I discuss below might have something to do with specificity.

[24] Though neither *sentir* 'feel' nor *trainer* 'lie around' take the auxiliary *être* on the relevant interpretation.

[25] For both this and the example which follows an object interpretation for the question is also available. I have controlled for this in asking speakers' judgements by putting them into a context which forces the subject reading as in (i).

(i) A ton avis, que révèle le mieux [le redressement de
in your opinion, what reveals the best the re-establishment of
l'économie], les chiffres de chômage ou le taux de l'inflation?
the economy the figures of unemployment or the rate of the inflation
'In your view what best reveals the economic recovery, the unemployment figures or the rate of inflation?'

(48) ? Que vous demanderait un voeu de célibat
what you would ask a vow of celibacy
'What would require a vow of celibacy from you?'

(49) ? Que réclame toute notre attention?
what demands all our attention
'What demands our full attention?'

What seems particularly relevant here is that in all these cases, on a subject interpretation,[26] 'what' seems to mean something like 'what particular thing'. In other words, *que* is being interpreted here as 'D-linked' to use the terminology of Pesetsky (1987), or if Kiss (1993) is right in equating the two, a specific or familiar indefinite. It is well known that many languages bar indefinites from occurring in the [Spec,IP] position, or require that they receive a particular type of interpretation either as a specific or a generic. In some languages (Modern Standard Arabic is one), the addition of a modifier may be sufficient to render the indefinite specific enough to be able to occupy this position. Clearly, some indefinites may appear in subject position in French but it may be that *que/quoi* are so resistant to a specific interpretation that, except where no other interpretation is available, as in an echo, it is rejected in [Spec,IP]. This ideas seems to be borne out by the contrast mentioned to me by Paul Hirschbühler (p.c.) between the multiple interrogation in (50) and the more complex one in (14) repeated here as (51).

(50) ?? Quoi trainait où?
what lay around where?
'What was lying around where?'

---

[26] (47) and (48) are open to object interpretations too; perhaps the fact that the object interpretation is more prominent in (i) than in (47) accounts for the fact that fewer speakers accepted it.

(i) Que démontre que l'économie se redresse?
what shows that the economy is re-establishing itself
'What shows that the economy is recovering?'

(51) ? Qui a dit que quoi trainait où?
who has said that what lay around where
'Who said that what was lying around where?'[27]

In (51) the context provides strongly for an interpretation in which the answer(s) to 'what' must be selected from a previously delimited set, much as is the case with 'which X' in English, which has been claimed to be associated with a necessarily D-linked interpretation. Of course, to determine whether this explanation is really on the right track much more detailed informant work would be required. However, the fact that many speakers will accept *quoi* as a subject on an echo interpretation is further suggestive of this view, since these are clearly specific. In addition, the fact that long-distance questions where *que* can escape [Spec,IP] are possible lends strong support to this view. Further, questions with an expletive subject, where *que* does not need to transit through [Spec,IP], are correctly predicted to be good under the Minimal Approach since when [Spec,IP] is filled by a non-wh-element, just as in object or adjunct questions the Wh Criterion cannot be satisfied without subsequent movement.[28]

Finally, whether it is ultimately correct to regard periphrastic questions like (52) as genuinely long-distance or not, they clearly differ from simple questions in their propositional force, which in many cases is a diagnostic of specificity. Thus in both English and French, (52) but not (53) presupposes that something did indeed happen.

---

[27] The ambiguity which appears in the English gloss if the complementiser is omitted here is not a factor in the French where embedded finite complementisers may be omitted only in interrogative clauses. The alternative interpretation of the English gloss would have to be rendered as in (i).

(i) Qui a dit ce qui trainait où?
who has said it that lay around where
'Who said what was lying around where?'

[28] These questions do suggest, however, that seeing strong features as categorial requirements only cannot be quite right. If it were, one would wonder why an expletive could not satisfy the requirement. This leads us back to a more traditional approach in which the element to be checked against the strong feature must bear compatible wh-features.

(52) Qu'est ce qui s'est passé?
what is it that is happened
'What was it that happened?'

(53) Que s'est-il passé
what is-it happened
'What happened?'

There remains work to be done on fleshing out the idea presented here but I am aware of only one problem with it. Pesetsky (1987) claims that elements like 'what the hell' are strongly non-D-linked. However, some speakers have been found to accept the following.

(54) Que diable te faisait imaginer que je serais chez moi à
what devil you made imagine that I would be house-my at
cette heure-là?
that hour-there
'What on earth made you think I'd be home at that time of day?'

I leave the resolution of this problem to further research.

4. Conclusion

In this paper we have seen that French questions possess a number of peculiarities which have major implications for our understanding of Wh Movement and how it is to be motivated within current syntactic theory. I have proposed a number of revisions to Rizzi's approach to questions to bring it into line with current thinking arguing in line with Chomsky (forthcoming) that checking is a one-way mechanism, at least with respect to wh-features. I have argued that the revisions proposed to Rizzi's theory help us to explain in part the restrictions on *que* questions which have been so widely discussed in the literature on French syntax. These revisions alone do not suffice, however; there is a further constraint on the position of *que* which I have proposed is a strongly non-specific indefinite barred from terminating in [Spec,IP]. The impossibility of *quoi* subject questions is thus accounted for without a requirement that subject question-words move and is perfectly compatible with a *Minimal Approach* to Wh Movement, *contra* Rizzi

(1991). The impossibility of *que* subject questions, on the other hand is attributed to economy considerations but their equivalents with expletive subjects are correctly predicted to be possible. Rather than invalidating the Minimal Approach then, French 'what' questions actually lend support to it.

REFERENCES

Aoun, J. and Y-h A. Li (1993) *Wh*-elements in situ: syntax or LF?, *Linguistic Inquiry* 24.199-238.

Baker, C.L. (1970) Notes on the description of English questions: the role of an abstract question morpheme, *Foundations of Language* 6.197-219.

Belletti, A. and L. Rizzi (1995) (eds.) *Parameters and Functional Heads: Essays in Comparative Syntax* , Oxford: OUP.

Chomsky, N. (1993) A minimalist program for linguistic theory. In Hale, K and J. Keyser (eds.) *The view from Building 20*, 1-52, Cambridge: MIT Press.

Chomsky, N. (forthcoming) Categories and transformations, Chapter 4 of *The Minimalist Program*, Cambridge: MIT Press.

Chomsky, N. and H. Lasnik (1977) Filters and control, *Linguistic Inquiry* 8.3

Friedemann, M-A (1991) *Propos sur la Montée du Verb en C^0 dans Certaines Interrogatives Françaises*, Mémoire de Licence, Université de Genève.

Goldsmith, J. (1978) Que, c'est quoi? que, c'est QUOI, *Recherches linguistiques à Montreal*, 1-13.

Goldsmith, J. (1981) Complementizers and root sentences, *Linguistic Inquiry* 12.554-573.

Higginbotham, J and R. May (1981) Questions, quantifiers and crossing, *The Linguistic Review* 1.41-80.

Hirschbühler, P. (1978) *The Syntax and Semantics of Wh Constructions*, Bloomington, Indiana: IULC.

Kayne, R. (1975) *French Syntax: The Transformational Cycle*, Cambridge: MIT Press.

Kayne, R. (1976) French relative *que*. In Lujan, M and F. Hensey (eds.) *Current Studies in Romance Linguistics*, 255-299, Washington: Georgetown University Press.

Kiss, K. (1993) Wh-Movement and specificity, *Natural Language and Linguistic Theory* 11.85-120.

Koopman, H. (1982) Theoretical implications of the distribution of *quoi*, *Proceedings of the North Eastern Linguistics Society*, 153-62, Amherst, MA: GLSA.

Léard, J-M. (1982) Essai d'explication de quelques redoublements en syntaxe du québécois; l'interrogatif-indéfini, *Revue québécoise de linguistique* 2, Montreal: UQAM.

Lefebvre, C. (1982a) (ed.) *La Syntaxe Comparée du Français Standard et Populaire*, 2 vols., Québec: Office de la Langue Française.

Lefebvre, C. (1982b) *Qui qui vient* ou *Qui vient*: voilà la question, in Lefebvre, C. (1982a).

May, R. (1985) *Logical Form: Its Structure and Derivation*, Cambridge: MIT Press.

McDaniel, D. (1989) Partial and multiple Wh-Movement, *Natural Language and Linguistic Theory* 7.565-604.

Obenauer, H-G. (1976) *Etudes de syntaxe interrogative du français*. Quoi, combien, *et le complémenteur*, Tübingen: Niemeyer.

Obenauer, H-G. (1977) Syntaxe et interprétation: *que* interrogatif, *Le Français Moderne* 45.305-341.

Pesetsky, D. (1987) Wh-in-situ-: movement and unselective binding. In Reuland, E. and A. ter Meulen (eds.) *The Representation of (In)definiteness*, 89-129, Cambridge: MIT Press.

Plunkett, B. (1993) *Subjects and Specifier Positions*, University of Massachusetts doctoral dissertation, Michigan: University Microfilms (also to be published by GLSA: Amherst).

Plunkett, B. (1994) The minimal approach to Wh Movement, ms.. University of York.

Sportiche, D. (1994) Clitic constructions, ms. UCLA.

Rizzi, L. (1989) *Relativized Minimality*, Cambridge: MIT Press.

Rizzi, L. (1991) *Residual Verb Second and the Wh-Criterion*, Technical Reports in Formal and Computational Linguistics, no. 2., Faculté des Lettres, Université de Genève republished in Belletti and Rizzi (1995).

Stroik, T. (1995) Some remarks on superiority effects, *Lingua* 95.239-258.

Williams, E. (1986) A reassignment of the functions of LF, *Linguistic Inquiry* 17.265-299.

EVENT STRUCTURE AND THE *BA* CONSTRUCTION*

Catrin Siân Rhys

University of Ulster at Jordanstown

1. Introduction

The controversy surrounding the *ba* construction within Chinese linguistics concerns the semantic content of *ba* and its relation to the matrix verb. On the one hand, it is argued to be a full lexical preposition, independently assigning a thematic role to its complement (Li 1985, Cheng 1986). On the other hand, it is claimed to be a dummy Case marker with no semantic content, inserted to license the direct object of the verb (Huang 1982, Goodall 1987). Constraints on *ba* and the interaction of *ba* with more general syntactic constraints in Chinese have the effect that the well formedness of *ba* fronting ranges from obligatory through preferred and optional to ill-formed. In its simplest form, however, the *ba* construction is an optional mechanism for fronting the object of a transitive verb:

(1) a. ta sha le fuqin.
he kill ASP father.
He killed his father.

b. ta *ba* fuqin sha le.
he father kill ASP.
He killed his father.

---

* Author's address: School of Behavioural and Communication Sciences, University of Ulster at Jordanstown, Newtonabbey, Co Antrim BT37 0QB.

York Papers in Linguistics 17 (1996) 299-332
© Catrin Siân Rhys

Under early assumptions in GB, the conclusion that the *ba* object was moved also forced the conclusion that *ba* itself was a semantically empty dummy Case marker inserted at S-structure, because of the Theta Criterion. Previous analyses have therefore tended to concentrate on the properties of the movement operation and the contexts in which it was obligatory.

With the advent of theories of functional heads, *ba* can be viewed as a base generated functional head with independent semantic properties but crucially no thematic grid. The constraints on the licensing of the *ba* construction then move to centre stage, as the properties of the functional head and its complement are determined. This is the approach taken in this paper. *Ba* is given a novel analysis in which it interacts with the thematic structure of matrix verb via a system of thematic mediation, but more importantly, it interacts with event structure via the hierarchy of aspectual roles proposed in Grimshaw (1990). This dual interaction allows us to capture both the formal aspects of *ba*, that have lead to its treatment as a dummy Case marker, and the interpretive effects of *ba*, which have lead to its analysis as a thematic head. Furthermore, I show that the analysis developed here has some interesting results for the argument structure of the *ba* construction, in addition to the desired effect of accounting for the relation between an affectedness constraint on the DP following *ba*, and the aspectual restrictions on the verb phrase in the *ba* construction.

Before investigating the constraints on the licensing of *ba*, the structure assumed for the *ba* construction is outlined along with some motivating data.

2. What is the structure of the *ba* construction?

The first observation to be made about the *ba* construction is that the apparent object of *ba* canonically gets its thematic role from the verb and appears in the post verbal complement position, as shown in the simple *ba* construction given in (1b) which relates to the canonical order in (1a) (repeated here):

(1) a. ta sha le fuqin.
she kill ASP father
She killed her father.

b. ta ba fuqin sha le.
she *ba* father kill ASP
She killed her father.

This suggests that *ba* is not a thematic role assigner and that the apparent object of *ba* is not a complement of *ba*, or at least is not assigned a thematic role by *ba*. This suggestion is strengthened by the observation that *ba* and its apparent object do not behave as a constituent with respect to movement. The following examples show that they cannot appear either postverbally, or sentence initially, or outside VP.[1]

(2) a. *ying lin sha le ba muqin.
Ying Lin kill ASP *ba* mother

b. *ba muqin ying lin sha le.
Ba mother Ying Lin kill ASP

c. *ying lin ba muqin zuotian yong dao shasi le.
Ying Lin *ba* mother yesterday use knife kill ASP

It should be noticed in this context that the apparent object of *ba* is licensed to appear in all the above positions without *ba*. It can also even appear in the preverbal *ba* position without *ba*, which suggests that in addition to not being a thematic role assigner, *ba* is not simply an inserted Case assigner.[2]

---

[1] See Y. H. A. Li (1985: 373) for more detailed argumentation that *ba* occupies a position within VP.

[2] Although of course an alternative interpretation of this fact is that when the object does appear in the *ba* position without *ba*, there is a null Case assigner, carrying the focus interpretation of the construction. However, the question of Case assignment in Chinese is not one I wish to address in this paper (see Rhys 1992). It has also been pointed out to me by a reviewer that it is not clear that the unmarked preverbal object is in fact in the same position as *ba*, since interaction with adverbials points to the unmarked preverbal object being outside VP.

If *ba* and its apparent object do not form a constituent, what, then, is the constituent structure involved? An important observation in this case is that *ba* imposes aspectual restrictions on the VP that follows it. So the following example is ruled out because the VP is stative and not perfective as required by *ba*.[3]

(3) *wo ba ta ai.
I *ba* her love

This relationship of *ba* to the VP, and the fact that it does not assign a thematic role to its apparent object, point to a structure in which the actual complement of *ba* is in fact the VP. Indeed *ba* does appear to behave like other functional heads that have a VP complement, in that the position of *ba* is fixed, as shown in (2), and iteration of *ba* is not licensed. Hence in the following example, either object of the double object verb *jiao* 'spray' can be *ba* fronted, but not both:

(4) a. ta ba hua jiao le shui.
he *ba* flowers spray ASP water
He sprayed the flowers with water.

b. ta ba shui jiao le hua.
he *ba* water spray ASP flowers
He sprayed the water on the flowers.

c. *ta ba hua ba shui jiao le.
he *ba* flowers spray ASP water
He sprayed the flowers with water.

In addition, reduplication of *ba* in the A-not-A structure, as in (5), shows that it is a verbal head in the verbal projection since only verbs can be negated by the negative particle *bu* that appears in the A-not-A reduplication:[4]

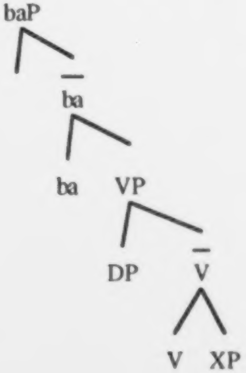
---

[3] This is a simplification of the aspectual restrictions as will become clear below.

(5) ni ba bu ba shu gei ta?
you *ba* not *ba* book give her

The evidence thus points to the following structure in which *ba* is a functional head with a VP complement. The apparent object then appears in the specifier of the VP complement governed by *ba*, but not theta marked by *ba*.[5] Henceforth this DP will be referred to as the *ba* DP, and *not* the *ba* object.

(6)



The relation between *ba* and the *ba* DP, is taken to be one of thematic mediation (see Rhys 1992 for motivation for such an analysis). The idea of thematic mediation comes from Grimshaw's discussion of the role of the prepositions *to* and *of* in licensing the arguments of nominals (Grimshaw 1990: 71). This idea is developed in Adger and Rhys (forthcoming), in which lexical heads have both argument structure and thematic structure and the Generalised Theta Criterion requires that thematic roles be assigned to arguments. In this approach, a thematic mediator is a functional head with argument structure but no thematic structure, which licenses a thematic role from a lexical head which either has no argument structure (e.g. nominals), or has an argument saturated by something other than the thematic role (e.g. nominal gerunds). It is this relationship of thematic mediation (and the a-role structure of *ba* to be discussed below) that gives the appearance of constituenthood to *ba* plus the *ba* DP, and yields the adjacency requirement of *ba* and the following VP, ruling out certain kinds of typical VP behaviour, e.g. coordination, VP-initial adverbs, etc.

---

[4] It has been pointed out by a reviewer that prepositions such as *gen* 'with' might also arguably be negated by *bu*. In Rhys 1992, however, I have argued that precisely this set of putative prepositions are in fact also verbal functional heads interacting with the thematic structure of the matrix verb.

[5] Note that this rules out adoption of any simple view of the VP internal subject hypothesis of Koopman and Sportiche 1991. For discussion of this see Rhys 1992.

3. Aspect and the constraints on *ba* fronting

With the exception of Cheng (1986), early accounts (e.g. Huang 1982) have concentrated on the structural properties of *ba* , and the contexts in which it is obligatory. The constraints on *ba* fronting have been assumed to be peripheral; a matter of semantics or even pragmatics. These accounts have therefore not attempted to explain the ungrammaticality of examples such as:

(7) * wo ba yige qianbao shi le.
I *ba* a purse find ASP

(8) * wo ba ta ai.
I *ba* her love

(9) * wo ba ji kanjian le.
I *ba* chicken saw ASP

(10) * wo ba qian you.
I *ba* money have

The unacceptability of (7) relates to the definiteness of the *ba* DP, which is generally claimed to be necessarily definite, but in this example is marked as indefinite by the indefinite article *yige*. The problem in (8) is one of aspect: *ba* fronting is not licensed when the verb constellation is stative. Both (9) and (10) are generally explained in terms of an affectedness restriction on *ba* DP, although (10) also does not meet the aspectual constraints on *ba* since the verb *you* 'have' is clearly stative.

Ba also interacts with the Postverbal Constraint (Huang 1982), the syntactic constraint on word order that makes object fronting obligatory when another constituent, whether complement or adjunct, appears in the postverbal position:

(11) a. wo ba ta mian le zhi.
I *ba* him cancel *le* job
I fired him.

b. *wo mian le zhi ta.
I cancel *le* job him

c. *wo mian le ta zhi.
I cancel *le* him job

Thus *ba* fronting may be obligatory (under the Postverbal Constraint), optional (in the simple *ba* construction as in (1)), ungrammatical (with certain aspectual classes), or preferred (in the resultative constructions to be discussed below).

Earlier GB accounts have generally acknowledged these descriptive generalisations about the *ba* construction but have taken the constraints on *ba* to be outwith the scope of a syntactic account. In the case of the definiteness restriction, it is certainly the case that this restriction is not specifically a property of the *ba* construction. Firstly, it is a more general property of word order in Chinese that preverbal NPs have a definite or specific interpretation whereas postverbal NPs have an indefinite interpretation. Thus in the case of ergative verbs where the subject is licensed either preverbally or postverbally, the difference in interpretation between the two subject positions is one of definiteness (examples from Sybesma 1992):

(12) a. tankeche lai le.
tanks come le
The tanks have come.

b. lai tankeche le.
come tanks le
There are some tanks coming.

It might also be argued that this definiteness restriction is the effect of the communicative function of *ba* which is to mark the object as 'given' information (Li 1971).[6] The aspectual restrictions and the affectedness restriction, on the other hand, should form an integral part of the analysis of *ba* licensing. Furthermore these two types of restrictions intrinsically interact. Cheng (1986) also acknowledges a connection between the notion of affectedness and the aspectual structure of the verb phrase. In her account, however, there is nothing inherent in either restriction from which this connection is derived. It is simply stated in terms of feature cooccurrence. Other than Sybesma (1992) whose analysis is discussed below, the only attempts to capture the affectedness restriction (Huang 1991, Cheng 1986) assume that there is a theta role <Affected Theme>.

In this paper, I suggest that the affectedness condition is not the consequence of a thematic role <Affected Theme>, nor is it a subclass of the thematic role <Theme>. Instead, based on Grimshaw (1990), I propose that it derives from an independent hierarchy of semantic roles distinct from thematic roles. Furthermore this second hierarchy is derived from the aspectual structure of the verb constellation. The interaction of the two restrictions on *ba* therefore derives from this relationship between the semantic hierarchy and aspectual structure.

---

[6] A reviewer has pointed out that the definiteness effects in the *ba* construction appear to be much more robust than for other preverbal DPs, and that the explanation for this may well lie in event structure of the *ba* construction, which would fit well with the general approach developed here.

3.1. Aspectual classes and an aspectual ontology

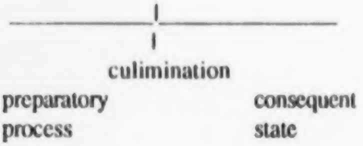
Since Vendler (1967), it has been generally acknowledged that the classification of predicates into aspectual classes accounts for their different behaviour with respect to temporal adverbials and aspect markers. Dowty (1979) details a number of diagnostics for determining aspectual class, and shows that the aspectual class of a clause can be influenced by the arguments of a verb as well as by the verbal constellation. Examples of the four aspectual classes given by Vendler and Dowty are as follows:

• *state*	know, love, be tall
• *activity*	run, walk, drive a car
• *accomplishment:*	kill, paint a picture, build a house
• *achievement*	recognise, reach, die

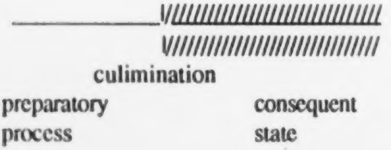
States relate to the traditional stative/non-stative distinction, a distinction which is maintained between states and the other classes, so that the general term for an aspectual class is *eventuality*, reserving the term *event* for the non-stative aspectual classes. Among the events, accomplishments and achievements differ from activities in that they have an *inherent* endpoint, a property often termed *telicity*. This telic/atelic distinction leads to a distinction in past tense aspects between *completion* and *termination* (Smith 1991). A telic verb with its inherent endpoint typically involves completion: the event *John ran to the shops* ends when John reaches the shops. An activity, an atelic verb with no inherent endpoint, simply terminates: *John ran*. Activities and accomplishments differ from achievements in that they involve duration.

Moens and Steedman (1988) develop an ontology of events based on the event structure template of (13) (over) which gives the internal structure of an event. Their proposal is that the different aspectual classes map differently onto this template. The telic property of accomplishments and achievements, mentioned above, is captured by a mapping involving both the culmination and consequent state, the difference between them being that the accomplishment also involves a

(13)

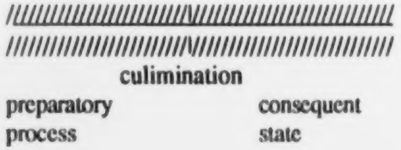

culimination
preparatory
process
consequent
state

(14)


culimination
preparatory
process
consequent
state

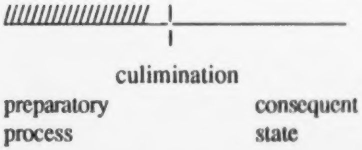
preparatory process. Hence, the achievement ***reach the top*** maps as in (14), where the event involves the culmination, i.e. reaching the top, and the consequent state of being at the top. Whereas the accomplishment ***build a house*** involves the preparatory process of building, in addition to the culmination, the completion of building, and the consequent state, the existence of the house, as in (15).

(15)


culimination
preparatory
process
consequent
state

An activity such as ***run***, on the other hand, involves neither culmination nor consequent state, but just the preparatory process part of the template:

(16)



The difference between *termination* and *completion* can now be reformulated as the difference between an event which culminates (completion) and an event that ends before culmination (termination). Moens and Steedman add an additional event to the traditional three; the *punctual* event. This is an instantaneous event which involves only a culmination and neither preparatory process, nor consequent state, for example *sneeze*.

The relationship between the subevents in this template, Moens and Steedman argue, is neither directly temporal nor causal (as proposed in Dowty 1979). Rather they show that it is a relation of *contingency*. In the analysis below, Moens and Steedman's system is adopted as it renders the internal structure of an event transparent, and offers a straightforward approach to the compositional building up of an event.

3.2. Grimshaw's aspectual roles

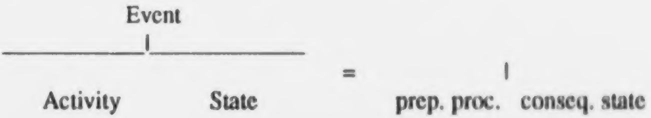
Grimshaw (1990), in an account of psychological predicates, suggests that there is a dimension of semantic analysis independent from thematic structure which is essentially causal in nature. The two classes of psychological predicates are represented by *frighten* and *fear* which have the same thematic analysis but are distinguished along this dimension: *frighten* is causative whereas *fear* is stative. The importance of this for Grimshaw is that it provides insight into the argument realisation of the two verb classes. In particular, it sheds light on the question of why, in the *frighten* class of predicates, the Theme is realised as the subject despite being lower on the thematic hierarchy. This fact now falls under the broader generalisation that *cause* arguments of causative predicates are always subjects. The causal status of arguments is thus indicative of an independent dimension of prominence relations that is distinct and autonomous from the thematic dimension:

(17) (Cause(other( )))

It is the alignment (or misalignment) of arguments across the thematic dimension and this causal dimension that yields differing behaviour in relation to argument realisation.

The contentful notion of *cause*, however, is too narrow. Neither agentive predicates, nor unergative predicates, nor psychological predicates show any of the effects of the misalignment of the two semantic dimensions, so their subjects must have some property in common which qualifies them for maximal prominence on the causal dimension. They are not however causatives. How then is this second dimension defined? Grimshaw suggests that the answer lies in the event structure of the predicates and that the dimension is aspectual in nature. Adopting a Vendler/Dowty approach to event structure, Grimshaw suggests that aspectual prominence derives from participation in the subevents of a complex event. For example, an accomplishment such as *break* is a complex event which breaks down into an activity and a state, which in Moens and Steedman's terms, are the preparatory process and the consequent state. (The Dowty/Vendler system does not separate the consequent state from the culmination.)

(18)



Under such an analysis, the cause argument is always associated with the first subevent, the preparatory process. Grimshaw generalises this to the claim that the argument that participates *only* in the first subevent of a complex event is aspectually more prominent than an argument that is associated with both or only the second subevent. I shall continue to refer to the aspectual role (a-role) assigned to that argument as <Cse>, although it should be understood that the causal interpretation stems not from the a-role itself but from the contingency relation between the two subevents of the complex event, i.e. it is in some sense epiphenomenal.

3.2.1. Aspectual roles in Chinese

Is there any evidence for this independent aspectual hierarchy in Chinese? The causal interpretation of (19) suggests that there is:

(19) wo ba tade chuangkou da-po le.
I *ba* her window hit-broken ASP
I broke her window.

The verb complex in this example, *da-po*, is a resultative compound formed from the two verbs *da* and *po*. The verb *da* means 'hit' and has as its core theta roles Agent and Theme, neither of which has a causal interpretation:

(20) wo da le tade chuangkou.
I hit *le* her window
I hit her window.

The verb *po* is an intransitive verb roughly translating as 'broken', with the single theta role Theme:

(21) tade chuangkou po le.
her window broken *le*
Her window is broken.

If we assume that the thematic structure of the compound *da-po* 'break' derives from the thematic structure of its two component verbs, then the overall thematic structure of the compound will be <Agent, Theme>, that is identical to the thematic structure of *da* 'hit', where the Theme of *da* 'hit' has identified with the Theme of *po* 'broken'. The compound, however, has a causative interpretation that is absent from either of the component verbs. This suggests that the interpretation of the subject of the compound as a Cause cannot be thematic. Turning to the event structure, on the other hand, we find that the compound is an overt realisation of the preparatory process-consequent state structure, in which the Agent is a participant of only the preparatory process, hence is assigned Grimshaw's a-role, <Cse>. Note that the *object* in (19) has an *affected* interpretation that is similarly absent in (20) and (21). This suggests that affectedness should also not be analysed as a property of the thematic grid as Huang and Cheng have both assumed, but derives from the aspectual dimension. This is the hypothesis addressed in the next section.

3.3. Affectedness, the aspectual dimension and *ba*

The first step in the hypothesis is to look to event structure for a participant that will be interpreted as affected. If this is the case then as well as the a-role <Cse>, we can define a second a-role <Aff>, and the aspectual hierarchy will be specified as:

(22) (Cause(Aff))

Consider the predicate *kill* in the sentence: *John killed the cat.* Here *John* is the <Cse> and *the cat* receives an interpretation as the affected object. If we turn now to the event structure of the predicate, we find that it is an accomplishment comprising a preparatory process, killing, and a consequent state, being dead. In particular we find that while *John* is the participant only of the preparatory process, and hence is assigned the a-role <Cse>, *the cat* is the sole participant of the consequent state. This points to a definition of the a-role <Aff> as the participant of a consequent state. If we look now at the Chinese translation of 'kill' the same appears to be true.

(23) Zhangsan sha le xiaomao.
Zhangsan kill ASP cat.
Zhangsan killed the cat.

Assuming that *sha* has the same lexical event structure as its English translation, *Zhangsan* is the Agent of the preparatory process and *xiaomao* is the participant in the the consequent state. Thus, we find again that the notions of *cause* and *affected* correlate with these roles in the event structure. We can, therefore, abstract away from the contentful notions of Cause and Affected and work in terms of aspectual subevents and their associated participants. Under this approach, we can now reformulate the affectedness constraint on *ba* in terms of event structure and aspectual roles. More precisely the *ba* DP can be viewed as the participant of a consequent state in a complex event. Thus the object of (23) can appear as a *ba* DP, whereas this is not possible with a verb such as *ai* 'love' that is a state and not a complex event:

(24) Zhangsan *ba* xiaomao sha le.
Zhangsan *ba* cat kill ASP
Zhangsan killed the cat.

(25) *Zhangsan *ba* xiaomao ai.
Zhangsan *ba* cat love

This seems to be a step in the right direction because it does look as though event structure rather than a contentful role is what is relevant. So in the following example, the object could not be said to be affected in any way, and yet *ba* fronting is licensed:

(26) ta ba yaoshi diu-le.
he *ba* key lose ASP
He lost the key.

The claim that *ba* picks out the participant of the consequent state in a complex event entails that a verb like *diu* 'lose' must be argued to be a complex event, having a consequent state, 'lost', that is predicated of the *ba* DP. Evidence for this comes from adverbial modification. If (26) is modified by an adverb of duration *sange xiaoshi* 'for three hours', the only interpretation available is that the consequent state of the key being lost lasted for three hours:

(27) ta ba yaoshi diu-le sange xiaoshi.
he *ba* key lose ASP three hours
He lost the key for three hours.

In fact, a comparison between the verbs that do allow *ba* fronting with the ones that do not, indicates that the feature that distinguishes the verbs that allow *ba* fronting is that their event structure involves a consequent state when the verb is combined with the aspect marker *le* (*le* is ambiguous between termination and completion). Examples are verbs such as *chi* 'eat', *xi* 'wash', *si* 'tear up', *wang* 'forget', *pian* 'cheat'. The verbs that do not allow *ba* fronting on the other hand all seem to be either states such as *renshi* 'know', or atelic processes such as *ting* 'listen', which either do not perfectivise (in the case of states) or involve only termination where the perfective *le* is licensed. The following are examples of verbs that do not generally license *ba* fronting: *tui* 'push', *shang* 'go up', *dai* 'carry', *xihuan* 'like'.

3.4. V-V compounds, consequent states and *ba*

The idea that *ba* picks out the participant of the consequent state of a complex event is supported by data from V-V compounds. There are two kinds of V-V compounds, conjunctive and resultative (Li 1990). The conjunctive ones are like *bangzhu*, where both halves of the compound mean *help*. They are all either punctual or processes, and do not break down into subevents. The resultative compounds are like overt realisations of the preparatory process--consequent state structure of the lexical complex events. So for example, *chi-guang* 'eat-empty' involves the process of eating and the consequent state in which the bowl is empty, and *chi-bao* 'eat-full' involves the process of eating and the consequent state of the eater being full:

(28) wo chi guang le fan.
I ate empty ASP rice
I ate up all the rice.

(29) wo chi bao le fan.
I ate full ASP rice
I ate rice and ended up full.

If *ba* picks out the participant of the consequent state, then we would expect *ba* fronting of the object to be licensed with *chi-guang* 'eat-empty', where the consequent state is predicated of the object *fan*, and not with *chi-bao* 'eat-full', where the consequent state is predicated of the matrix subject. This expectation turns out to be correct:

(30) wo ba fan chi-guang le.
I *ba* food eat-empty ASP
I ate up all the rice.

(31) *wo ba fan chi-bao le.
I *ba* food eat-full ASP

Thus we can explain why it is that where the interpretation of the V-V compound is ambiguous, as with *qi-lei* 'ride tired', *ba* fronting is licensed, but yields only the interpretation where *lei* 'tired' is predicated of the object:

(32) a. wo qi lei le neipi ma.
I ride tired *le* that horse
either: I rode that horse and it got tired.
or: I rode that horse and got tired (myself).

but

b. wo *ba* neipi ma qi-lei le.
I rode that horse and got it tired.

3.5. Aspectual role assignment and functional heads

So far it is claimed that the *ba* DP occupies a particular position in the event structure of the clause. This is implemented using Grimshaw's notion of an aspectual hierarchy. In particular, the *ba* object must realise the second most prominent role in the aspectual hierarchy, i.e. <Aff>. Furthermore, this information must be part of the syntactic representation of the *ba* construction. So how can *ba* be specified to pick up the second role in an aspectual structure? Recall that *ba* is claimed to be a thematic mediator, parallel to the analysis of the coverbs given in Rhys (1992). It is thus a functional head with a VP complement, licensing the thematic roles from its VP complement via its own argument structure. Given this structure, I propose that *ba* actually assigns both <Cse> and <Aff>; <Aff> to the DP in the specifier position of its VP complement, and <Cse> to its own specifier. In other words, by analogy with thematic roles, it has the a-role structure (Cse(Aff)).

In fact, I will adopt the strong claim that a-roles are not assigned at all by lexical heads but only by functional heads such as *ba*. Thus the ambiguity in example (32) (repeated here) arises because no a-roles are assigned:

(32) wo qi lei le neipi ma.
I ride tired le that horse
either: I rode that horse and got tired.
or: I rode that horse and it got tired.

Since no a-roles are assigned here, neither DP is explicitly marked as the participant of the consequent state. When *ba* is projected, it assigns the a-role Aff which explicitly marks the *ba* DP as the participant in the consequent state. Assuming the requirement of the standard Theta Criterion that all arguments must be assigned a thematic role, a-role assignment is not sufficient to satisfy the Theta Criterion, so the *ba* DP has to receive its thematic role from a lexical head. This explains the conflict between the apparent semantic content of *ba*, and the evidence that the *ba* DP receives its thematic role from the verb. *Ba* does have independent semantic content but it is aspectual and not thematic. Effectively what *ba* does, then, is assign aspectual prominence relations, which interact with the event structure of its complement. In other words, by virtue of the a-roles that it assigns, *ba* requires that the event structure of its complement VP be a complex event.

This is somewhat different from Grimshaw's approach in that a-roles here are syntactically and not lexically assigned. In Grimshaw's approach aspectual prominence relations are a lexical feature on an argument derived from the lexical representation of the event structure of a lexical head. In the Chinese data that we are considering here, the event structure of the predicate is not lexical, but rather is built up compositionally as part of the syntax. A-roles therefore cannot be part of the lexical representation of the thematic role assigning head. In fact, even in Grimshaw's system it transpires that the representation of the aspectual structure cannot simply be projected from the lexical semantic representation of the individual predicate, but involves the projection of an abstract event structure template that breaks down into two subevents: an activity and a state or change of state. Aspectual prominence is determined on the basis of participation in this abstract event template. The difference between the two approaches thus reduces to the level at which the template applies.

Under this analysis we now have an explanation for the following difference in interpretation between a sentence with the object in canonical postverbal position and the corresponding *ba* construction, observed by Sybesma (1992).

(33) wo qi lei le neipi ma.
I ride tired ASP that horse
I rode that horse and it got tired.

(34) wo ba neipi ma qi lei le.
I *ba* that horse ride tired ASP
I rode that horse and *got it* tired.

The difference between the two sentences relates to causativity in that there is a stronger causal interpretation in the sentence involving *ba* fronting. Recall that the relationship between subevents in the Moens and Steedman template is one of contingency. The semantics of the resultative compound, however, further specifies the relationship as one of causation. In example (33), we therefore have a relation of causation between the preparatory process of riding, and the consequent state of being tired. However, no a-roles are assigned and the causation is interpreted as a relation between events. In (34), on the other hand, the a-roles are explicitly assigned and the causation is relation between the participants of the subevents, since the subject is marked as the Agent of the causation, the Cse, as well as the thematic Agent, and the *ba* DP is marked as the Aff. In this way, explicit assignment of the a-roles in a causal complex event will yield a stronger causal interpretation.

4. V-V compounds and argument structure

Whether in the V-V compound the consequent state is predicated of the subject or the object of the process or is ambiguous is not a linguistic issue; it is world knowledge not syntax that tells us that in example (29) rice cannot be full. The fact that the consequent state has to be predicated of one of the arguments of the first subevent is however a matter of syntax. Li (1990) suggests that it is Case restrictions that force argument identification. However, this fails to account for the restrictions on licensing (see the discussion in Rhys 1992). Assuming, however, that identification has somehow been forced, the extension of Grimshaw's system developed here gives us the argument structure of the V-V compound. So, for the V-V compound *qi-lei* 'ride-tired', one interpretation is that the horse being ridden ends up tired, in other words, the Theme of *ride* identifies with the experiencer of *tired*. I will represent this as follows, where the indexes attached to the thematic roles refer to the subevents that the arguments participate in, i.e. 1 is the preparatory process, and 2 is the consequent state:

(35) qi lei
Ag-1, Th-Exp-1+2

This means that the Agent is higher in the aspectual structure than the Theme, because it participates only in the preparatory process. In other words, in terms of the aspectual hierarchy (Cse(Aff)), the Agent is compatible with the <Cse> role. The Th-Exp then is the participant of the consequent state and can be assigned the a-role <Aff>. We thus capture the fact that *ba* fronting of the object is licensed under this interpretation.

So what about the alternative interpretation where the Agent identifies with the Experiencer?

(36) qi lei
Ag-Exp-1+2, Th-1

Reading the aspectual prominence relations directly from the indices assigned to the thematic roles, we find that the change in interpretation also yields the reverse aspectual prominence relations. It is the Theme that participates only in the preparatory process, whereas the Agent is identified with the Experiencer and so participates in both subevents. The <Aff> aspectual role therefore cannot be assigned to the Theme, which is now highest on the aspectual rating. The fact that *ba* fronting of the object is not available for this interpretation is thus captured. However, Grimshaw's system for assigning aspectual prominence also predicts that the Theme should be licensed as subject since it is only associated with the first subevent, and the specification of *ba* predicts that the Agent-Exp should be licensed as a *ba* object. This is because it is indexed as the participant of the consequent state and therefore should satisfy the a-role <Aff>. This prediction holds and the following example is acceptable:

(37) ma *ba* wo qi lei le.
horse *ba* I ride tired ASP
The horse tired me out riding it.

In fact, this arrangement of thematic and aspectual relations yields precisely the set of examples which Sybesma calls the *causative ba* sentences.

(38) Zhei-jian shi ba Zhang San ku-lei le.
This-CL case *ba* Zhang San cry-tired ASP
This thing got Zhang San tired from crying.

(39) ku-lei
Ag-Exp-1+2, Th-1

In fact, under this system we also get some explanation for the ergativity shift phenomenon that Sybesma discusses. Sybesma argues that the *ba* construction involves an abstract CAUS predicate which gets phonological content either by V raising or by insertion of *ba* which he claims is a dummy element. An important feature of his analysis is the claim that the complement of this abstract CAUS predicate is ergative. Adopting Hoekstra's (1988) account of resultatives, Sybesma essentially claims that the resultative V-V compounds involve at D-structure a matrix verb with a resultative complement and *assumes* that the resultative complement triggers a shift to ergativity in the matrix verb, suppressing the external argument of the matrix verb. The test for ergativity in Chinese is the postverbal subject. Hence, while ***ku*** 'cry' does not license its subject postverbally in (40), in the resultative compound ***ku-lei*** 'cry-tired', he claims it does:

(40) *ku le yixie hao ren.
cry ASP some good people
(intended: Some good people cried.)

(41) ku-lei le yixie hao ren.
cry-tired ASP some good people
Some good people cried themselves tired.

Similarly:

(42) ku shi le shoujuan.
cry wet ASP handkerchief
The handkerchief got wet from crying.

Under my system, it is no surprise that such examples are ergative. In the mapping from aspectual structure to argument structure, Grimshaw argues that ergative/unergative distinction relates to whether the single argument predicate maps onto the first or second subevent of the event template. A single argument predicate that maps on to the first subevent, the preparatory process, will be unergative, whereas the single argument predicate that maps onto the second subevent, the consequent state, will be ergative. In fact, exactly what this predicts for (41) is not clear, since it maps on to both subevents and the single argument is associated with both subevents. This is reflected in native speaker judgements, which are divided over whether (42) necessarily involves an implicit Cause argument, in which case, the predicate is not ergative but transitive. In (42) on the other hand, the predictions are clear. Since the only argument expressed is associated with only the consequent state, it will be licensed as the internal argument and the overall predicate will be ergative.

5. Resultative complements

This analysis also carries over to the phrasal resultative using the particle *de*. In this construction a consequent state is expressed by a clause in complement position introduced by *de*, which is cliticised onto the matrix verb:

(43) ta qi de ma hen lei.
she ride *de* horse very tired
She rode so much the horse got tired.

(44) ta qi de hen lei.
she ride *de* very tired
She rode so much she got tired.

In the examples above, there is no matrix object competing with the resultative complement. Where the matrix object is expressed in this construction, fronting of the object is obligatory, by the Postverbal Constraint, as the resultative complement saturates the postverbal complement position. However, the fronted object can be licensed preverbally either by *ba* or by verb reduplication, and the different licensing mechanisms trigger different interpretations. Adopting Huang's (1991) insight that these resultative constructions are, at some level of representation, complex predicates, they are assigned a complex event structure parallel to the lexically formed V-V compounds. Again licensing by *ba* forces the reading where the *ba* DP is the participant of the consequent state. Compare:

(45) wo ba ma qi de lei le.
I *ba* horse ride *de* tired ASP
I rode the horse and got it tired.

(46) wo qi ma qi de lei le.
I ride horse ride *de* tired ASP
I rode the horse and got tired.

The reason that the resultative construction is important to the study of *ba* is that *ba* fronting of the subject of the resultative complement is licensed even where the DP in question is clearly an argument only of the embedded clause and not of the matrix clause:

(47) wo ku de Zhangsan hen shangxin.
I cry *de* Zhangsan very sad
I cried so much that Zhangsan was very sad.

(48) wo ba Zhangsan ku de hen shangxin.
I *ba* Zhangsan cry *de* very sad
I cried so much that Zhangsan was very sad.

The matrix verb in these sentences is *ku* 'cry' which on its own does not license an object, either in canonical object position or as a *ba* DP:

(49) *wo ku le Zhangsan.
I cry ASP Zhangsan

(50) *wo ba Zhangsan ku le.
I *ba* Zhangsan cry ASP

The *ba* DP must therefore be theta marked in the embedded clause. This is a property *only* of resultative complements; other embedded clauses do not permit *ba* fronting of their subjects. While this is problematic to explain for purely syntactic accounts of *ba*, these facts simply fall out from the aspectual account of *ba* that I have developed here.

In general there is, for every V-V compound, a corresponding resultative construction. However, there is a difference in interpretation between the V-V compound and the resultative construction relating to causality. In the same way that *ba* fronting in a V-V compound yields a stronger causative interpretation than the non-*ba* fronted form, so the resultative compound has a stronger causative interpretation than its V-V compound counterpart:

(51) a. wo qi lei le neipi ma.
I ride tired *le* that horse
I rode the horse and it got tired.

b. wo qi de neipi ma lei le.
I ride *de* that horse tired *le*
I rode that horse and got it tired.

The particle *de* thus clearly does have some semantic content. In particular, it has a similar semantic effect to *ba*. In the following analysis I adopt Huang's basic intuition that the resultative construction forms a complex predicate with the matrix verb, but I argue that this is a property of the event structure and not syntactic as Huang assumes. A detailed analysis of *de* resultatives is however beyond the scope of this investigation. What we are interested in here is the interaction of the resultative complement with *ba* and with the event structure of the sentence.

5.1. Resultative *de* and event structure

The basic claim here is that *de* is a functional head which combines with its complement and with the matrix clause to form a complex event. More precisely, there is, as part of the semantic representation of *de*, a rule that essentially means that *de* combines two independent events, to yield one complex event. Using bracketing to mark subevents this can be represented as shown:

(52) (e1) de (e2) → (E(e1)(e2))

This captures Huang's intuition that these are complex predicates without forcing unmotivated abstraction in the syntax. Under this analysis, it is a complex predicate in that it yields a single complex event. This interaction of *de* with event structure is reflected syntactically in that *de* is also an a-role assigner assigning the two a-roles (Cse (Aff)). In fact, it may be possible to derive the rule in (52) from the a-role structure of *de*. It assigns the a-role <Aff> to the DP that it governs in the subject position of the resultative clause, and assigns the most prominent a-role <Cse> to the subject of the matrix clause.[7] If both *de* and *ba* are projected, the a-roles are forced to identify as they map onto to the same complex event. The only difference in interpretation is one of causality; there is a stronger causal interpretation when both functional heads are projected. This, as we have seen, can be attributed to the relationship between causality and the a-roles assigned. Apart from this, the following have the same interpretation:

---

[7] Note that I am only claiming an aspectual parallel between *de* and *ba*. Hence, we would not necessarily expect parallel behaviours in other

(53) a. Zhangsan ku de Lisi hen shangxin.
Zhangsan cry *de* Lisi very sad
Zhangsan got Lisi sad with his crying.

b. Zhangsan ba Lisi ku de hen shangxin.
Zhangsan *ba* Lisi cry *de* very sad
Zhangsan got Lisi sad with his crying.

These two have the same interpretation because the DPs in question are assigned the same a-roles. This suggests an explanation for the following, otherwise confusing, observation. Where the matrix verb has both a transitive and an intransitive reading but there is no matrix object, the matrix verb is nonetheless interpreted transitively and the subject of the resultative is necessarily interpreted as the matrix object:

(54) Zhejian shi jidong de Zhangsan ku le.
This matter excite *de* Zhangsan cry *le*
This matter excited Zhangsan so much that he cried.
not: This matter was so exciting that Zhangsan cried.

---

respects. For example, a reviewer has pointed out that while the *ba* DP must be overt, the DP following *de* can be empty. There are a couple of potential sources for this difference. Huang 1984 shows that empty complements are in fact instances of wh-movement, whereas empty subjects can be pro. Furthermore, only *ba* is a thematic mediator. So essentially, the question seems to boil down to why a thematically mediated argument cannot be wh-moved. Note that this is true for all the coverbs which I have argued should be analysed as thematic mediators in Rhys 1992.

As is seen from the translation, although the matrix verb *jidong* 'excite' appears to be used intransitively, it must be interpreted transitively with the meaning *excited Zhangsan*. This can be understood as the effect of the a-role assigned to Zhangsan, which is canonically realised as an object. It also explains the marked preference for the corresponding *ba* fronted sentence.

This analysis in terms of a-roles explains both the object interpretation of the subject of the resultative and the availability of *ba* fronting. It also captures the parallel causality effects of the resultative complements and *ba* fronting in the V-V compounds.

6. Why do we need to refer to the internal structure of the event?

Until now, we have been referring to the internal structure of an event. However, the eventuality involved in the *ba* structures we have addressed so far is always an accomplishment with a fixed internal structure. If this is the case, then do we really need to build so much structure into the analysis? Or could the analysis simply make reference to the aspectual category of accomplishment, rather than the consequent state in a complex event? For example, one could imagine an analysis in terms of the object of an accomplishment formed by a simplex, or complex predicate.

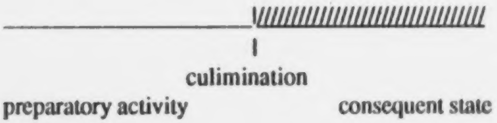
One response to the criticism that the account is building more structure than is necessary might be to point to other linguistic phenomena that require reference to the internal structure of the event. Grimshaw's work on argument structure in English discussed above, for example, requires reference to the internal structure of the event via an event template. Stronger motivation, however, comes from the *ba* construction itself. In the following data, examples are given in which the *ba* construction is licensed, but the eventuality involved is clearly not an accomplishment. Such data would obviously cause problems for an analysis in terms of accomplishment. However, the internal structure of the event does involve a consequent state as expected under this analysis.

6.1. Inchoatives

A frequently observed counterexample to the claim that *ba* is only licensed in accomplishments is the following:

(55) wo ba ta ai shang le.
I *ba* her love PRT ASP
I fell in love with her.

The aspectual classification of such an utterance is inchoative, where inchoatives are thought to pick out the begining part of the event. What then is the internal structure of an inchoative? Going back to the Moens, and Steedman template, inchoatives are also analysed as involving a culmination and consequent state.



The difference between the accomplishment and the inchoative is that the culmination in the inchoative marks the *initial* bound of the event, whereas in the accomplishment it marks the final bound (Moens p.c., Kamp p.c., Dowty 1979). Thus, in an example such as (55), the culmination is the falling in love and the consequent state is the being in love. We can show that the consequent state is indeed part of the linguistic representation of 'fall in love' by the contradiction in (56), where the entailed consequent state is negated:

(56) ! I fell in love with her but I never loved her.

Thus the inchoative is clearly shown to involve a consequent state, which would lead us to expect that *ba* fronting with inchoatives is licensed.

6.2. Progressive - *zhe*

Another apparent counterexample to the descriptive restriction of *ba* to bounded events is the use of *ba* with the progressive marker *zhe*.

(57) ta ba yifu bao-zhe.
he *ba* clothes bundle-PROG
He is bundling up the clothes.

At first blush, such an example appears to be an irredeemable problem for the account of *ba* given here. However, appearances can be deceptive and in this instance, it is the translation of *zhe* as a progressive, that leads to the deception. In fact a much more appropriate translation would be as a resultative along the lines of 'He has the clothes bundled up' with the resultative particle 'up'. In fact, Carlota Smith argues very convincingly that 'in its basic meaning *-zhe* is a resultative stative' (Smith 1994: 122).

The common representation of *zhe* as a progressive stems from its additional use as a backgrounding particle, in examples such as the following:

(58) Xiao Li zuo zhe kan shu.
Xiao Li sit *zhe* read book
Xiao Li is reading sitting down.

In this use *zhe* loses the resultative interpretation, and has a simple activity reading with no internal structure at all. If the analysis of *ba* given here is correct, we would predict then that *ba* fronting with the backgrounding use of *zhe* is not licensed. And indeed, the data in (59) shows that this is the case:

(59) *Xiao Li ba yifu bao zhe chang ge.
Xiao Li *ba* clothes bundle *zhe* sing song.

Thus again we find that it is the specification of *consequent state* that is crucial to the distribution of *ba*.

6.3. Directionals

An additional interesting result arises with examples such as the following from Wang (1987):[8]

(60) ta zhengzai ba chuan wang shui li tui
she now *ba* boat towards water in push.
She's pushing the boat into the water.

It is generally assumed to be the case since Vendler (1967) that an activity verb with a goal yields an accomplishment, e.g. *run to the park*, whereas an activity verb with a directional adverb or complement remains an activity, and this can be tested for using Dowty's time adverbial tests, where *in*-adverbials are appropriate with accomplishments but not with activities. Hence:

(61) a. Michelle drove to the university in five minutes flat.
b. ?Michelle drove towards the university in five minutes flat.

Activity verbs with directionals are not, however, straightforward activities, hence the oddness of (62a) as compared to (62b):

(62) a. ?Michelle drove towards the university for five minutes
b. Michelle drove around the university for five minutes.

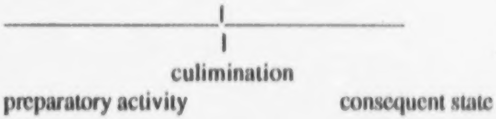
(62a) is by no means ill-formed but does seem to require some contextual explanation, hence the improvement in (63):

---

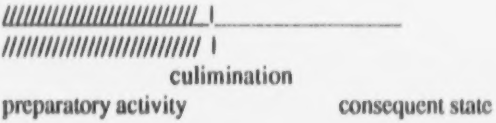
[8] Note that this example provides counterevidence to the common assumption that *ba* fronting is not licensed with monosyllabic verbs, based on examples such as the following:
(a) *wo ba ni sha.
I *ba* you kill.
This is judged as unacceptable, but becomes acceptable combined with the aspectual particle *le*. This not, in fact, a question of syllabicity, but rather of event semantics, since the same expression is licensed in a conditional:
(b) ruguo wo ba ni sha, ...
If I *ba* you kill, ...
Thus, the explanation for (a) will be in terms of event semantics and compatible with the approach to *ba* developed here.

(63) Michelle drove towards the university for five minutes before changing her mind and turning back.

We can begin to get a handle on the difference between the simple activity in (62b) and the activity plus directional in (62a), by referring again to Moens and Steedman's event template:



The simple activity in (62b) involves just the first part of the template, the activity part, and terminates, but has no culmination, as follows:



The activity plus directional also refers to the activity part of the template, but in addition it provides information about the consequent state that would be reached if the event culminated rather than simply terminating. That is, although a presupposition of (62a) is that Michelle does not end up at the university, it is also true to say that part of the meaning of (62a) is that if the activity of Michelle driving towards the university does not terminate, then there is an inherent culmination point, the arrival at the university, and the consequent state of being at the university. In other words, the consequent state is not entailed but can be inferred, and clearly must be part of the representation of a directional expression.

Accounting for (60), therefore means that we must extend the analysis of *ba* to incorporate not just consequent states that are entailed by the event structure but also ones that can be logically inferred. This might seem like an undesirable weakening of the initial analysis. However, closer examination of the aspectual classes in Chinese suggests that this is necessary to account for simple lexical accomplishments.

The question of the existence of lexical accomplishments in Chinese is controversial. Based on the following examples, Tai (1984) and Heinz (1984) both argue that in Chinese there is no grammaticalisation of telicity; that is that the culmination and consequent state that are the defining features of accomplishments are not part of the lexical meaning of verbs such as *sha* 'kill'.[9]

(64) wo sha le ta liang ci dou mei si.
I kill ASP her 2 times all not die
I tried to kill her twice but she didn't die.

(65) Zhangsan xue-le Fawen, keshi mai xue-hui.
Zhangsan learn *le* French but not learn-able
Zhangsan studied French but never learnt it.

(66) wo mai le sanben shu, keshi mei mai-dao.
I buy *le* three books, but not buy-arrive
I tried to buy three books but didn't manage to.

Smith (1990) argues that these verbs are telic but that the perfective particle *le* in Chinese does not have the same interpretation as perfective in a language such as English, but is ambiguous between termination (no culmination) and completion (culmination). An alternative approach which avoids the disjunctive analysis of *le* is to argue that the aspectual structure of a lexical accomplishment in Chinese does include a culmination and a consequent state but that the consequent state is not an entailment of the verb and hence is defeasible. The relevance of this problem here is that *ba* fronting is licensed showing that the consequent state required by *ba* need not be an entailment of the predicate:

---

[9] Native speaker judgements on these examples vary enormously. They are give here in order of decreasing acceptability with only the first being universally accepted.

(67) wo ba ta sha le liang ci dou mei si.
I *ba* her kill ASP 2 times all not die
I tried to kill her twice but she didn't die.

Returning to the example in (60), there would seem then to be independent motivation that a consequent state that is *inferrable* from the directional expression is sufficient to license *ba*.

7. Conclusion

Much of the earlier controversy around *ba* stems from dissension over whether or not *ba* has any independent semantic content. Either *ba* was assumed to be a purely formal particle, the function of which was to assign Case, or it was argued to have semantic content and this was assumed to translate into thematic content. Under the hypothesis that abstract Case does not play a role in Chinese (Rhys 1992), *ba* cannot be a Case marker. However, I have also argued against the second option of assuming thematic content to *ba*. Instead I have argued for a second kind of semantic information that plays a role in syntactic description; namely event structure. I have shown in this paper that the affected interpretation of the *ba* DP is the consequence, not of a particular thematic role, but of the a-role assigned by *ba*. In this way, the constraints on *ba* are captured and shown to be intrinsically linked, and the supposed control facts of Huang (1991) fall out. Furthermore the relationship between *ba* and causality is now understood as a consequence of the contingency relations between subevents of a complex event. The extension developed here of Grimshaw's theory of the interaction between aspectual structure and thematic structure and the consequences for argument structure was shown to predict both the ergativity shift in certain V-V compounds, and the well-formedness of the causative *ba* sentences.

Thus this paper provides further evidence for a model of syntax in which there is considerable interaction between the syntactic representation and the level of event structure, cf. Ramchand (1993), McClure (1994).

REFERENCES

Adger, D. and Rhys, C. S. (forthcoming) Eliminating disjunction in lexical specification. In P. Coopmans, M. Everaert and J. Grimshaw (eds.) *Lexical Specification and Lexical Insertion*. Hillsdale, N-J: Lawrence Erlbaum Assoc.

Cheng, L. L. S. (1986) *Clause Structures in Mandarin Chinese*, MA Thesis, University of Toronto.

Dowty, D. (1979) *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Grimshaw, J. (1990) *Argument Structure*. Cambridge, Mass.: MIT Press.

Heinz, M. (1984) Chinese: A language without lexical accomplishment, Ms, University of Wisconsin - Madison.

Hoekstra, T. (1988) Small clause results. *Lingua* 74.101-139

Huang, C-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, PhD thesis, MIT

Huang, C-T. J. (1984) On the distribution and reference of empty pronouns. *LI* 15.531-574

Huang, C-T. J. (1991) *Complex predicates in control*, Ms University of California at Irvine.

Koopman, H. and Sportiche, D. (1991) The position of subjects. *Lingua* 85.211-258

Li, Y. F. (1990) On Chinese V-V compounds. *NLLT* 8.177-207

Li, Y. H. A. (1985) *Abstract Case in Chinese*, PhD thesis, University of Southern California

McClure, W. (1994) *Syntactic Projections of the Semantics of Aspect*, PhD thesis, Cornell University.

Moens, M. and Steedman, M. (1988) Temporal ontology and temporal reference, *CL* 14.15-28

Ramchand, G. (1993) *Aspect and Argument Structure in Modern Scottish Gaelic*, PhD thesis, Stanford University.

Rhys, C. S. (1992) *Functional Heads, and Thematic Role Assignment in Mandarin Chinese*, PhD thesis, University of Edinburgh

Smith, C. (1994) Aspectual viewpoint and situation type in Mandarin Chinese, *Journal of East Asian Linguistics* 3.107-146

Sybesma, R. (1992) *Causatives and Accomplishments: The case of Chinese ba*. Leiden: HIL.

Vendler, Z. (1967) *Linguistics in Philosophy*. Ithaca: Cornell University Press.

EXPLANATION OF SOUND CHANGE. HOW FAR HAVE WE COME AND WHERE ARE WE NOW?

Charles V. J. Russ

Department of Language and Linguistic Science
University of York

1. Introductory: The development of explanations

1.1 Extralinguistic explanation

Early explanations of sound change were often sought in extralinguistic factors such as the climate, or the physiology of the speakers. Thus, the second or High German sound shift in which the initial Germanic voiceless stops became affricates , e.g. p, t, k became [pf], [ts], [kx] (the velar only in Upper German). This change was carried through in initial position before vowels and, in the case of p and k before /l/ and /r/, while t was only shifted before /w/. This was viewed by some linguists as being caused by the Alpine climate. Since it was carried through most completely in Southern Germany, Austria and Switzerland, which are mountainous regions, it was assumed that there was a causal relationship between the sound shift and the climate or geography of the region. This view was advanced by serious linguists, but it was to be refuted by Jespersen. He pointed out that the tendency to affrication of voiceless stops was not confined to mountainous regions, but that there was a strong tendency to affricate initial prevocalic t in the colloquial speech of Copenhagen (Jespersen 1922: 256f). Similar explanations were given for the First Germanic Sound Shift (see survey in Russ 1978: 169-73).

Most scholars have been hesitant to explain sound changes in terms of extralinguistic factors, but the most widely accepted way that extralinguistic factors are used to explain change is in the substratum theory. The Latin of the Roman Empire was imposed on countries with

York Papers in Linguistics 17 (1996) 333-349

© Charles V. J. Russ

other native languages, e.g. Celtic in France, and consequently the natives of these countries imposed the features of their own language on the Latin they learned. These original, or substrate languages died out in most cases, but have left their mark in the way Latin has developed in different countries. For instance some linguists claim that the French change of Latin û to [y:], e.g. Latin *mûrus*, French *mur*, is due to the Celtic substrate, or that the shift of f to h, which is then lost in pronunciation in Spanish, e.g. Latin *facere*, Spanish *hacer* 'to do', is due to the Basque substrate. In general it is accepted that some changes may be due to substrate languages but the actual extent of this is not agreed (see Pellegrini 1980 for further references).

Much of the use of extralinguistic factors in explaining sound changes has been speculative and many changes have been found which could not be put down to these factors. Bloomfield, and structural American linguists in general, thought that the search for explanations or causes of sound change was fruitless. Bloomfield said explicitly 'The causes of sound change are unknown' (Bloomfield 1935: 385). Hockett (1958), for example, contains no references to the causes of sound change.

1.2 Internal linguistic explanations

Other linguists, notably the Prague group, swung away from extralinguistic causes completely to the other extreme, wanting to see the causes of linguistic change in the linguistic system itself. They, and later Martinet, are the prime exponents of this view. They did not regard sound laws as blind, as the Neogrammarians did, nor fortuitous as de Saussure (1916: 127) thought, but rather purposeful. Sound change was seen as teleological, goal directed. This might take various forms. There might be various 'goals', the removal of peripheral phonemes, e.g. /ɔɪ/ in English (Vachek 1964), or of phonemes with a low functional yield, e.g. the merger of /ɛ̃/ and /œ̃/ or /a/ and /ɑ/ in French (Martinet 1961: 210f), or the making of an asymmetrical system symmetrical. A persuasive example of the last type of change in Swiss German dialects has been given by Moulton (1961: 155-182). Classical Middle High German is assumed to have the following short vowel system:

i	ü	u
e	ö	o
ë		
ä		a

This is an asymmetrical system, since the back vowels have one less tongue height than the front unrounded vowels. In the North East of Switzerland this system was made symmetrical by the split of /o/ into /o/ and /ɔ/: 'The asymmetry of the Middle High German system lay in the fact that the front vowels contained one more relevant level than the back vowels. In the West and Centre this asymmetry was removed by decreasing the number of front vowels. In the North and East the asymmetry was removed by increasing the number of back vowels: the /o/ of Middle High German *ofen, hose* (New High German *Ofen* 'stove', *Hose* 'trousers') split into modern /ofə/ ≠ /hɔsə/' (Moulton ibid., 172f [Translation CR]). The result of this change was a symmetrical short vowel system. There was a complementary split of Middle High German /ö/ into /ö/ and /œ/. Jakobson attempted to illustrate his teleological view of sound change by applying it to Russian. For example, the *akanje*, the merging of unstressed ạ and ọ, in Russian and other dialects, is seen as resulting from the change of the correlation: musical accent - unstressed vowels, to expiratory accent - unstressed vowels (Jakobson 1971: 92ff).

Martinet, building on the work of the Prague school, developed the notion of the push-chain and the drag-chain. When a phoneme moves phonetically in one direction and approaches another phoneme, e.g. /A/ > /B/, then /B/ may also move towards another phoneme, /C/, /B/ > /C/. This chain reaction is a push-chain, /A/ pushes /B/ towards /C/. Another possibility would of course be that /A/ and /B/ merge, but Martinet is more interested in the cases where this does not happen. If, taking the three phonemes /A/ /B/ /C/, /C/ moves first, away from /B/, then /B/ may well also be dragged into the space vacated by /C/, and then /A/ may be dragged into the space left vacant by the shifting of /B/ (Martinet 1952: 5ff; 1955: 48ff). For instance, in early Old High German there were two dental obstruents (excluding the sibilants) /ð/, and /d/. The latter was shifted to /t/ and the space thus left vacant was then filled by the shift of /ð/ to /d/ (Penzl 1975: 86). This kind of chain reaction is called a drag-chain. This approach to sound change was taken up by many linguists, among them Weinrich, who, in his studies of Romance sound changes, sought to explain them without using extralinguistic factors (Weinrich 1958: 5ff).

This type of approach to sound change has been criticized on several grounds. The push-chains, drag-chains, development towards a symmetry are said to be only tendencies (King 1969: 191ff). There are asymmetrical sound systems - for instance many Upper German and Central German dialects have two front vowel phonemes /e/ and /ɛ/ but only one back vowel phoneme /o/. Enough evidence seems to have been produced that in certain cases sound changes can be explained in terms of other changes, but there are also many changes which cannot be thus explained. Also any teleological view of sound change is circular. In the Swiss German example taken from Moulton it could be seen that the result of the split of Middle High German /o/ into /o/ and /ɔ/ was a symmetrical short vowel system. The result and the cause are regarded in fact as being the same thing (Anttilla 1989: 193f). In other instances these explanations are only considered to be descriptions. This was the position taken up by a reviewer of Weinrich (1958): 'A mon avis, et j'espère pouvoir montrer par la suite qu'il est bien fondé, la phonologie diachronique ne pourra être que descriptive, ne saura jamais répondre à la question: POURQUOI? Pour répondre à cette question, il faut toujours recourir à des facteurs externes' (Togeby 1959/60: 402). However, although criticisms have been levelled against this approach, it has produced many results which have been accepted as worthwhile by many linguists.

1.3 Generative linguistics and explanation

The scepticism which Bloomfield expressed at ever finding explanations of sound changes was continued by generative grammarians. The most extreme position is that taken up by Postal: 'There is no more reason for languages to change than there is for automobiles to add fins one year and remove them the next, for jackets to have three buttons one year and two the next' (Postal 1968: 283). On the whole, the generative school has been criticized for not seeking explanations for sound change. This is not entirely fair, since opinions among generative linguists seem to vary. King, for instance, is not as sceptical as Postal: 'If there is little risk in being a cynic about the origin of phonological change, there is also very little profit. In fact linguistics has a great deal to lose by the position that the cause of phonological change is beyond principled research' (King 1969: 190f). However, he does not give any clear explanation of sound change. One approach to explanation in sound change can be illustrated from Kiparsky's historically orientated article entitled 'Explanation in phonology'. He states: 'I have suggested a way in which the concept of a 'tendency', which lends functionalist discussions their characteristic unsatisfactory fuzziness, can be made more precise in terms of hierarchies of optimality, which predict specific consequences for linguistic change, language acquisition, and universal grammar' (Kiparsky 1972: 224). For Kiparsky, explanation in sound change is determined by constraints such as the conservation of functional distinctions, e.g. a sound change will tend not to eliminate number or tense endings. When sound changes cause phonological alternation within an inflectional paradigm, e.g. lengthening of short vowels in open syllables, North German [taːgɜ], but nom. [tax] or [tak], the alternation will tend to be removed to make the paradigm regular, cf. standard German, *Tage*, *Tag*. Some sound changes may act together in a 'conspiracy' to produce a certain kind of phonological structure. However these constraints do not always apply. For instance modern German still retains the phonological alternation between medial voiced obstruents and final voiceless obstruents. This has been in existence since late Old High German and yet has not been levelled out except in a few dialects.

1.4 Some recent developments

Most textbooks on historical linguistics give surveys of some of the kinds of explanations and causes that have been outlined in 1.2 and 1.3, adding remarks on how sociolinguistics can help account for why particular variants are selected by a language (Anderson 1973: 3-5; Jeffers and Lehiste 1979: 88-105; Aitchison 1981: 111-69). A landmark in the discussion on explaining linguistic change is Lass (1980) who comes to the conclusion that to explain linguistic change must also entail predicting it. Therefore, since prediction of changes is impossible, explanation is also impossible. However, Lass's conclusion challenged many linguists to search for explanations. Vennemann (1983) says that he will continue explaining linguistic change, particular in terms of what is and what is not a possible change. Bennett (1983) argues that Lass sets too high a standard for explanations and that linguists should continue to search for them: 'The best way to be sure of not discovering the causes of linguistic change is to adopt the working assumption that there are no such causes. But if we seek, we may find' (1983: 20). Aitchison (1987) in a contribution to a workshop set up because of the impact of Lass's claim maintains that linguists should at least be able to sketch possible paths of development for changes. Lass (1987), himself, seems to offer a less pessimistic scenario, urging linguists to take a more long-term view of changes in languages in any attempts at explanation. Kiparsky (1988) as well as surveying different types of change and causes expresses the view that the linguist should not be surprised or despair if one language develops a structure in one way whereas another language develops the same structure in a different way. This balancing act of using both internal, functional explanations as well as external, sociolinguistic ones is continued in recent works (Hock 1986: 627-61, and 1992: 228-31; Crowley 1992: 191-203; Ohala 1994: 4050-55). McMahon (1994: 46) expresses the problem by saying 'We shall consider further, generally particularistic and non-predictive, explanations of changes in all components of the grammar, while striving to find general causes and motivations for change.' The wish to find causes and the conviction that they may be discovered is thus very much alive.

2. Types of explanatory statement

We have so far used the term 'explanation' without any real definition. In the following sections four ways in which it is used will be examined and their usefulness evaluated. Much of this, paradoxically, derives from a little known review by Bloomfield (1934).

2.1 General Historical Explanation

Bloomfield (1934: 34f) outlines this type of explanation in the following terms: 'Where the facts are accessible, we can define a feature of a language in terms of some earlier habit plus a change of habit'. This is a general form of explanation: something in the present can always be explained by saying that it represents something in the past plus a change. The strange shape of a house, for example, may be explained historically by saying that in the past there were two houses, which were than joined together. A linguistic example would be the explanation that umlaut in New High German is due to the fact that in Old High German the vowels affected were followed by an i̲, î̲, or j̲: 'Umlaut is used to express the change from a, o, u and au to ä, ö, ü and äu respectively ... . The cause of these vowel-changes can, as a rule, not be seen in modern German: in order to understand them, one requires to go back to the earlier stages of the language' (Eggeling 1961: 348). This type of explanation is not restricted to linguistics but it is common to all disciplines which have a historical branch. It has also fallen out of favour since it mixes the synchronic and the diachronic. De Saussure in his discussion of the necessity of separating the synchronic from the diachronic uses umlaut of noun plurals as part of his argument. He takes two stages in the development of German and English: At stage A the plural of some nouns is formed by adding -i̲: Old High German *gast*, *gasti*, OE *fôt*, *fôti*. At a later stage B, the plural is formed by changing the vowel, and in the case of German, adding -e̲: *Gast*, *Gäste*, *foot*, *feet*. For de Saussure, these ways of marking the plural have no historical connection. The only connection is between individual forms, e.g. *gasti*, which becomes *Gäste* (de Saussure 1916: 120ff). For him, umlaut in New High German would not be explicable in terms of Old High German. This attitude of de Saussure's seems to have influenced linguists in turning away from the diachronic study of language. This represents, in other disciplines as well as linguistics, 'a general loss of faith in the efficacy of historical explanation. We try to understand our present position by analysing the component forces in play, not by tracing *post facto* the long chain of major forces which have brought it about but may have ceased to operate' (Trim 1959: 19). This type of explanation is too unrestricted to account for why sound changes proceed along one particular path in one language but along a different path in another.

2.2 Universals of Sound Change

Another approach is to look at the universal nature of some sound changes. Some similar patterns occur in different languages. For instance, the raising of long and mid vowels has not only caused diphthongization in English, but also in Dutch, and probably also in German (Lass 1976). There is not an infinite number of sound changes but a restricted number. If these can be characterized, then an explanation can be attempted for a much smaller number. For the Neogrammarians, sound laws were fixed to one place and one dialect at one time. Consequently they did not believe in universals of sound change. For them, what was universal was that sound laws had no exceptions. However the whole question of universals has been discussed not only on a synchronic level but also on a diachronic level. This has chiefly taken the form of characterizing the possible forms of linguistic change and to what constraints they are subject (Kiparsky 1972; Vennemann 1982: 149-54; Labov 1994). Universals can help to explain sound changes in that they reduce the number of possible sound changes to a finite number. A sound change is deemed to have been 'explained' if it is assigned to a more general process. Sound change is viewed as consisting of a set of meta-rules: palatalization, nasalization and so on, from which a language selects one, which, subject to certain language specific constraints, will proceed in a defined way. For instance, if a language palatalizes consonants, first the velars will be affected, then the dentals and finally the labials. It will not affect labials only, or dentals only. The consonants (only obstruents have so far been considered) will be palatalized before high front vowels first, then before mid front vowels and finally before low vowels (Chen 1973). As an example, Italian has palatalized Latin **k** only before front high and mid vowels: Latin *civitatum*, *centum*, Italian *città*, *cento*, but this has not occurred before low vowels: Latin *cantare*, Italian *cantare*. French, on the other hand, has palatalized Latin **k** before **a** as well: French *cité*, *cent*, *chanter*. This approach does not completely solve the problem of causation of linguistic change, but it does attempt to overcome the *ad hoc* explanation of individual changes. Thus the change of Latin **k** to [tʃ] and further to [ʃ] in French is not seen as an isolated change but as part of the larger change of palatalization. Chen cites examples from many different languages which make his thesis seem plausible, but he has to admit that there are exceptions. In Ancient Greek IE /kw/ and /t/ are palatalized to /t/ and /s/ respectively before /i/ and /e/. According to Chen's scheme, if a dental stop has been palatalized then a velar stop will have been palatalized as well. The reason for this exception, he says, is that IE /kw/ and /t/ are involved in a drag-chain. IE /s/ became /h/ in Ancient Greek, initially and medially, and the space left by the shifting of medial IE /s/ was filled by the palatalization of IE /t/ before /i/ in certain cases (there are exceptions to this).[1] The gap created by the change of /t/ to /s/ before /i/ was then filled by IE /kw/ becoming /t/ before /i/ and /e/.[2] Language specific changes like this drag-chain in Ancient Greek can invalidate the universal trend of palatalization. This may well turn out to be an isolated case, but on the other hand it belies the strong predictive power that Chen would like his theory to have.

Another approach to the problem of universals has been to set up universal strength hierarchies. For example, if obstruents are deleted or subject to lenition in a language, velars are most likely to be deleted first, then dentals and finally labials (Foley 1977: 28). Lass and Anderson (1973: 183-87), in their study of Old English obstruents, come to a different conclusion. When stops become weakened to fricatives the order is: dentals first, then labials and finally velars. Certain kinds of statements as to what are natural classes differ sometimes according to the language or period of the language concerned. This search for universal hierarchies is still very speculative and more detailed studies must be made available before it can be proved to have a more solid foundation. A phenomenon which is similar to strength hierarchies is the concept of the *Reihenschritt*.[3] If one phoneme of a phonetic order changes, then all the other phonemes of the same order change in the same way. A classic example is provided by the First Germanic Sound Shift where each member of each order of

---

[1] Buck 1933: para. 141: 'The assibilation of τ before ι is seen in large classes of words. But τ may also remain unchanged before ι, and the precise conditions governing this difference of treatment cannot be satisfactorily formulated.'

[2] Chen 1973 takes his interpretation from Allen 1957-8: 122f.

[3] Pfalz 1918 used *Reihenschritt* for vowel changes. A free translation in English might be 'parallel development'.

consonants changed its manner of articulation: the voiceless stops p, t, k became the voiceless fricatives f, θ, x, the voiced aspirated stops bh, dh, gh became either voiced stops or voiced fricatives according to their position in the word b/v, d/ð, g/ɣ, the voiced stops b, d, g became voiceless stops p, t, k (Fourquet 1954). Similarly all the Middle High German long high vowels (î, iu, û) diphthongized, not just one or two of them. The concept of *Reihenschritt* has been adopted by Martinet (1952: 17) to show how sound changes proceed by changes in distinctive features. In generative grammar the fact that parallel groups of sounds may change has been accounted for in terms of 'natural classes': 'Phonological changes tend to affect natural classes of sounds (p, t, k, high vowels, voiced stops), because rules that affect natural classes are simpler than rules that apply only to single segments' (King 1969: 122). The use of the word *tend* is significant in this quotation since these changes do not always take place. On the basis of natural classes one cannot always predict that of three voiceless stops, if t becomes an affricate, then p and k will become affricates as well. This may perhaps happen, as it does in some Upper German dialects, but it is by no means automatic.

Any universals that do exist seem, at the moment, to be only universal tendencies (even Chen 1973: 183 uses the term 'tendency'). Similar changes can be seen at work in many genetically unrelated and geographically widely dispersed languages. The important thing that this search for universals has shown is that sound change is not random but, all things being equal, sound changes, e.g. palatalization, will proceed in a predictable way, e.g. affecting velars first, then dentals and finally labials. But unfortunately in languages all things are not equal. Many other factors intervene. There may be the influence of the rest of the sound system, the morphology and syntax, and external influences from other dialects or languages. The social prestige of certain forms and their spelling may influence changes. All these factors may and do interfere in the smooth effectuation of these universal tendencies. There seems no way of predicting when these other factors will intervene. The search for universals has still not supplied an answer to the problem of the explanation of sound change in general.

2.3 The Predictive Power of Linguistic Explanation

This level of explanation can be characterized as the one 'in which we could account for the occurrence of a certain linguistic change at a certain place and time: e.g. Why did pre-Germanic change p, t, k to f, θ, h or why did English analogically extend the -s pl. of nouns? The answer would be a correlation of linguistic change with some other recognizable factor enabling us to predict the occurrence of a linguistic change whenever this factor was known' (Bloomfield 1934: 39f). Bloomfield sets this up as a goal to be reached, but does not offer, here or elsewhere, any solution. Nor, we must say, has any linguist to date. Chen, who deals with prediction in phonological change, has to set his sights lower: 'Even though we cannot predict that palatalization will take place in language X, we can nevertheless predict that if palatalization occurs at all it will spread along two dimensions or axes' (Chen 1973: 177). Once a sound change has taken place, its course can be predicted within certain limits, but we cannot predict why palatalization should take place in French but not in Dutch. This has been called the 'actuation problem' by some scholars: 'Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at different times?' (Weinreich, Labov and Herzog 1968: 102). For instance, why did the Germanic long high vowels diphthongize in German, English and Dutch but not in the Scandinavian languages? This type of question is the strongest and most interesting demand that could be made of a theory of explanation in historical linguistics. Unfortunately no answer can be given to it with the present state of linguistics, and it is doubtful whether there will ever be an answer.

2.4 The Explanation of Specific Changes

One of the most widespread interpretations of 'explanation' is the explaining of one event by another. Bloomfield puts this in the following way: 'A favoured earlier event, the 'cause', pulls a kind of invisible string which, in some metaphysical sense, forces the occurrence of a later event, the 'effect'' (Bloomfield 1934: 34). This assumes that one can connect some linguistic effects but not others. For instance, in the Germanic languages many original final vowels have been lost or reduced to [ə]. That is one linguistic event. It is also assumed that the stress accent in Germanic, instead of falling potentially on any syllable, became fixed on the root syllable. This represents another linguistic event. Most linguists link these two events together, the fixing of the stress accent causing the weakening and loss of unstressed syllables: 'The strong stress accent on the stem (or first syllable) *caused* in Germanic a progressive weakening of unaccented syllables' (Prokosch 1939: 133). Similarly the mutation of the long and short back vowels a̲, o̲, u̲ in the Germanic languages at various times has occurred before an i̲, î̲, or j̲ in the following syllable. In this case it is usually said, not that one event caused another, but that one factor, the existence and nature of the following i̲, î̲, and j̲, caused the change known as i̲-mutation or umlaut. The following explanation illustrated this clearly: 'There are two types of mutation in O.E., one A., which affects back vowels is caused by a following i or j, the other, B., which affects front vowels, is caused chiefly by u, or o, in some dialects also by a' (Wyld 1921: para. 103). This mode of explanation refers chiefly to individual conditioned changes. Where changes are not phonetically conditioned, the explanatory power of one change or factor in terms of another one is not so convincing. Attempts have been made to explain one unconditioned change in the light of another. This is the type of event which Martinet has dubbed push- or drag-chain. The Great Vowel Shift in English has been explained in this way. The two most important steps in the vowel shift are the diphthongization of the long high vowels ME î̲ and û̲, and the raising of the long mid vowels ME ê̲ and ô̲. Scholars have postulated causal relationships between these changes. Luick thought that the raising of the mid vowels happened first and caused the already existing high vowels to diphthongize, while Jespersen, on the other hand, thought that the diphthongization of ME long î̲, û̲ created a hole, into which the mid vowels ME ê̲, ô̲ were dragged (Lass 1976: 51-102; 1992).

It is very often not possible to establish with any accuracy the direction of the explanation in unconditioned changes such as this. Documentary evidence may be lacking or inconclusive. These explanations of changes in terms of other factors or events have one great drawback: they are not final explanations. It may be the case that the raising of the mid vowels caused the diphthongization of the high vowels, or, that the fixing of the stress accent on the root syllable caused the weakening or loss of unstressed vowels. Even so there still remains the question of why the mid vowels were raised in the first place, or why the stress in Germanic became fixed to the root syllable. In other words, final causation is not provided for at this level. The type of explanation discussed here is of a specific sound change or changes. These will probably only occur in one language or in related languages and be tied to a particular period in that language. Most linguists would accept that this level of explanation, linking events to other events, as cause and effect, is indeed possible but that it is a weak form of the explanation of sound change.

3. Conclusion

What can be reasonably demanded of a linguistic theory is that it should explain language specific changes. Other types of explanation are far more difficult, if not impossible, to formalize. Research into universals may help, but much more evidence for many more different processes will have to be forthcoming before it is based on a surer footing

Most linguists, however, are agreed that languages are subject to change and that there is variation in the spoken chain. Where they differ is on the emphasis placed upon this. The fact that language is subject to variation does not explain sound change (this variation is simply a characteristic of language), but it does point to the possible origin of sound change. Variation in the spoken chain produces variants in pronunciation, grammar and vocabulary. The important thing is what happens to these variants once they have arisen for whatever reason. Two things are important here. The variants may be idiosyncratic and not spread at all, or they may find their way into the linguistic system (Samuels 1972: 140). It is at this point that the question 'why?' may begin to be asked. Here we find ourselves at the level of *ad hoc* language specific explanations. These entail what has been called the 'transitional problem', i.e. what intermediate forms there are, and the 'embedding problem', i.e. how does a change fit into (a) the linguistic system as a whole, and (b) into the social structure of the users of the language concerned? There is also the 'evaluation problem', i.e. how the speakers themselves reacted to the change (Weinreich, Labov and Herzog 1968: 184ff). The question 'why?' seems only answerable in the case of why a particular variant was selected by the linguistic system in a certain case, rather than saying why one was not selected.

Explanations or causes of sound changes can be given as long as it is realized that they merely entail connecting phenomena to their effects, the reason for the selection of a particular variant or process may be due to several factors, in other words there may be multiple causation (Malkiel 1967). All such explanations are ad hoc, even though they represent a selection from a restricted range of sound changes (Samuels 1972: 155f). The ultimate causes of sound change are unknown but in many cases we can see with varying degrees of confidence what the immediate causes are.

REFERENCES

Aitchison, J. (1981) *Language Change: Progress or Decay*. London: Fontana.

Aitchison, J. (1987) The language lifegame: prediction, explanation and linguistic change. In W. Koopmann, F. van der Leek, O. Fischer and R. Eaton (eds.) *Explanation and Linguistic Change*. (Current Issues in Linguistic Theory 45). Amsterdam: Benjamins. 11-32.

Allen, W. S. (1957-8) Some problems of palatalization in Greek. *Lingua* 7.113-33.

Anderson, J. M. (1973) *Structural Aspects of Language Change*, London: Longman.

Anttila, R. (1989) *Historical and Comparative Linguistics*, Amsterdam, Benjamins. This is basically the same as the 1972 edition, *An Introduction to Historical and Comparative Linguistics*.

Bennett, P. (1983) The nature of explanation in historical linguistics. *York Papers in Linguistics* 10.5-22.

Bloomfield, L. (1934) Review of W. Havers, *Handbuch der erklärenden Syntax*. *Language* 4.32-40.

Bloomfield, L. (1935) *Language*. London: Allen & Unwin. American edition 1933.

Buck, C. D. (1933) *Comparative Grammar of Greek and Latin*. Chicago: University of Chicago Press.

Bynon, T. (1977) *Historical Linguistics*. Cambridge University Press.

Chen, M. (1973) Predictive power in phonological description. *Lingua* 33.171-91

Crowley, T. (1992) *An Introduction to Historical Linguistics*. Oxford University Press.

Eggeling, H. F. (1961) *Dictionary of Modern German Prose Usage*. Oxford: Clarendon Press.

Foley, J. (1977) *Foundations of Theoretical Phonology*. Cambridge University Press.

Fourquet, J. (1954) Die Nachwirkungen der ersten und zweiten Lautverschiebung. *Zeitschrift für Mundartforschung* 22.1-33.

Haudricourt, A. and A. Juilland (1949) *Essai pour une histoire structurale du phonétisme français*. 2nd edition. The Hague: Mouton.

Hock, H. H. (1986) *Principles of Historical Linguistics*. Berlin - New York: Mouton de Gruyter.

Hock, H. H. (1992) Causation in language change. In W. Bright (ed.). *International Encyclopedia of Linguistics*. Oxford University Press. Vol. 1. 228-31.

Hockett, C. (1958) *A Course in Modern Linguistics*. New York: Macmillan.

Jakobson, R. (1971) *Selected Writings I. Phonological Studies*. 2nd edition. The Hague: Mouton.

Jeffers, R. J. and I. Lehiste (1979) *Principles and Methods for Historical Linguistics*. Cambridge, Mass.: MIT Press.

Jespersen, O. (1922) *Language*. London: Allen & Unwin.

Keller, R. (1963) Zur Phonologie der hochalemannischen Mundart von Jestetten. *Phonetica* 10.51-79.

King, R. D. (1969) *Historical Linguistics and Generative Grammar*. Englewood Cliffs, New Jersey: Prentice Hall.

Kiparsky, P. (1972) Explanation in phonology. In S. Peters (ed.) *Goals of Linguistic Theory*. Englewood Cliffs, New Jersey: Prentice Hall.

Kiparsky, P. (1988) Phonological change. In F.J. Newmeyer (ed.) *Linguistics. The Cambridge Survey. Vol. I. Linguistic Theory. Foundations*. Cambridge University Press. 363-415.

Labov, W. (1994) *Principles of Linguistic Change I. Internal Factors*. Oxford: Blackwell.

Lass, R. (1974) Linguistic orthogenesis: Scots vowel quantity and the English length conspiracy. In C. Jones and J. M. Anderson (eds.) *First*

International Conference on Historical Linguistics. Amsterdam: North Holland. 311-43.

Lass, R. (1976) *English Phonology and Phonological Theory* .Cambridge University Press.

Lass, R. (1980) *On Explaining Language Change*. Cambridge University Press.

Lass, R. (1987) Language, speakers, history and drift. In W. Koopmann, F. van der Leek, O. Fischer and R. Eaton (eds.) *Explanation and Linguistic Change*. (Current Issuses in Linguistic Theory 45). Amsterdam: Benjamins. 151-76.

Lass, R. (1992) 'What, if anything, was the Great Vowel Shift?', in *History of Englishes. New Methods and Interpretations in Historical Linguistics*. M. Rissanen et al. (eds.), Berlin - New York: Mouton de Gruyter. 144-55.

Lass, R. and J. M. Anderson 1975) *Old English Phonology*. Cambridge University Press.

Malkiel, Y. (1967) Mutiple versus simple causation in linguistic change. In *To Honor Roman Jakobson*. The Hague: Mouton. Vol 2. 1228-46.

Martinet, A. (1952) Function, structure and sound change. *Word* 8.1-32.

Martinet, A. (1955) *Économie des changements phonétiques*. Berne: Francke.

Martinet, A. (1961) Élements de linguistique généerale. Paris: Armand Colin.

McMahon, A. M. S. (1994) *Understanding Language Change*. Cambridge University Press.

Moulton, W. G. (1961) Lautwandel durch innere Kausalität. *Zeitschrift für Mundartforschung* 28.227-51.

Ohala, J. (1994) Sound change. In R. E. Asher (ed.) *The Encyclopedia of Language and Linguistics*. Vol 8.4050-55.

Pellegrini, G. B. (1980) Substrata. In R. Posner and J. N. Green (eds.) *Trends in Romance Linguistics and Philology*. The Hague: Mouton.

Penzl, H. (1971) *Vom Urgermanischen zum Neuhochdeutschen*. Berlin: Schmidt.

Pfalz, A. (1918) Reihenschritte im Vokalismus. In *Beiträge zur Kunde der bayerisch-österreichischen Mundarten I. Sitzungsberichte der Kaiserlishen Akademie der Wissenschaften in Wien. Phil.-hist. Klasse* 190-2. Vienna: Hölder. 22-42.

Postal, P. (1968) *Aspects of Phonological Theory*. New York: Harper Row.

Prokosch, E. (1939) *A Comparative Germanic Grammar*. Philadelphia: Linguistic Society of America.

Russ, C. V. J. (1978) Kausalität und Lautwandel. *Leuvense Bijdragen* 67.169-82.

Samuels, M. L. (1972) *Linguistic Evolution*. Cambridge University Press.

Saussure, F. de (1916) *Cours de linguistique générale*. Ed. T. de Mauro. Critical edition. Paris: Payot.

Togeby, K. (1959-60) Les explications phonologiques historiques sont-elles possibles? *Romance Philology* 13.401-13.

Trim, J. L. M. (1959) Historical, descriptive and dynamic linguistics. *Language and Speech* 2.9-25.

Vachek, J. (1964) On peripheral phonemes of modern English. *Brno Studies in English* 4.7-100. Reprinted in *Selected Writings in English and General Linguistics*. The Hague: Mouton, 1976.

Vachek, J. (1970) Remarks on 'The Sound Pattern of English'. *Folia Linguistica* 4.24-31.

Vennemann, T. (1982) *Grundzüge der Sprachtheorie*. Tübingen: Niemeyer.

Vennemann, T. (1983) Causality in linguistic change. Theories of linguistic preferences as a basis for linguistic explanations. *Folia Linguistica Historica* 4.7-26.

Weinreich, U., Labov, W. and U. Herzog (1968) Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (eds.) *Directions for Historical Linguistics*. Austin: University of Texas Press.

Weinrich, H. (1958) *Phonologische Studien zur romanischen Sprachgeschichte*. Münster: Aschendorffsche Verlagsbuchhandlung.

Wyld, H. C. (1927) *A Short History of English*. Oxford: Blackwell.

HAS IT EVER BEEN 'PERFECT'? UNCOVERING THE GRAMMAR OF EARLY BLACK ENGLISH[*]

Sali Tagliamonte

Department of Language and Linguistic Science
University of York

1. Introduction

Genetic relationships between varieties are often assessed by cross-linguistic comparisons of the tense/aspect system. This is especially true of African American Vernacular English (AAVE), whose verbal delimitation paradigm has been the subject of intense study for decades. This is in part due to the ongoing and still contentious debate on whether its present system developed from a prior creole or directly from the vernacular British varieties spoken by early white plantation staff. The sheer complexity and abundance of grammatical apparatus concentrated in this area of the grammar make it an excellent site for examining the differences and similarities amongst related varieties.

Over the last few decades the frequently used domains of the verbal system have been extensively exploited. In the area of copula usage and past tense expression, the underlying systems of AAVE and other varieties of English were found to be similar, though AAVE tends to extend English rules through the application of additional phonological

---

[*] I gratefully acknowledge the generous support of the Social Sciences and Humanities Research Council of Canada, in the form of research grants #410-90-0336 and #410-95-0778 for the project of which this research forms part. I thank Salikoko Mufwene for his insightful comments and especially his help in making sense of semantic vs. morphological tense and aspect, Marjory Meechan for her meticulous critique of a final version of this manuscript and Shana Poplack for comments and encouragement all through.

York Papers in Linguistics 17 (1996) 351-396
© Sali Tagliamonte

and grammatical processes.[1] In other areas of the verb system, such as present time reference, the patterning of surface forms, although atypical of contemporary varieties of standard English, has been shown to constitute reflexes of linguistic change whose patterns of variability reflect the state of the English vernaculars to which the slaves were exposed (Poplack and Tagliamonte 1989, 1991),[2] while simultaneously differing from the behaviour proposed for creoles (e.g. Tagliamonte et al. 1996). But these findings have not been univocal. Some researchers such as Winford (1991), DeBose (1994), and DeBose and Faraclas (1994) claim that contemporary AAVE preserves traces of a creole grammar. Thus, despite decades of research, the origins of AAVE remain controversial.

One area of the tense/aspect system which presents a test in point for this issue is what I will refer to here as the PERFECT. In standard English the PERFECT is typically equated with the morphosyntactic construction *have* + past participle, as in (1).[3]

(1) AUXILIARY *HAVE* + PAST PARTICIPLE:
  a. Some of them ***have regretted*** it already. Yes, many of 'em ***have regret*** it already. (SE/006/171-173)[4]
  b. It ***been*** so long ***I've forgotten.*** (SE/020/87)
  c. I ***have been told*** that if they know you handling money, they raise your wages. (SE/010/1005-7)
  d. That was the first they learnt me and I'm old and it ***have remained*** here. (SE/002/115-6)

---

[1] See Baugh 1980; Fasold 1971, 1972; Labov 1969, 1972a; Labov et al. 1968; Pfaff 1971; Poplack and Sankoff 1987; Tagliamonte and Poplack 1988; Wolfram 1969, 1974.

[2] See also Poplack & Tagliamonte 1994 for the plural.

[3] In these data the main verb of the *have* + past participle construction can surface as a weak verb without inflection or as a strong verb with preterit morphology, in addition to the standard English past participle form, as illustrated in the second verb phrase in (1a).

[4] Codes in parentheses identify the speaker and line number in one of two corpora, Samaná English (SE) or the Ex-slave Recordings (ESR). For details of the corpora see below.

In AAVE the infrequency of verbal constructions with *have* coupled with the plethora of other forms used for comparable, though not entirely similar functions, e.g. auxiliary *be*, as in (2), preverbal *done*, as in (3), bare past participles, as in (4) and *ain't* + verb, as in (5), have been used as evidence of an underlying creole system.

(2) AUXILIARY *BE* + VERB:
  a. I'***m pass*** a lot of trouble. (SE/002/374)
  b. Now they have so many houses. They all ***is made*** it one thing. (SE/003/480-2)
  c. I'***m forgot*** all them things. (SE/015/257)
  d. Well, with me nothing ***is happen***, nothing strange. (SE/006/144)
  e. Let me see, I'***m*** near ***forgot*** what I was to holler. (ESR/001/43)

(3) PRE-VERBAL *DONE*:
  a. Plenty ***done gone*** and they'*s* lose their life. (SE/005/476)
  b. I ***done been*** to Miami, Hollywood ... (SE/010/1032)
  c. So much trouble ***done pass***. (SE/002/113-4)
  d. Grandpa was always saying them old oxens ***done run*** off in-***runned*** off in the river with us. (ESR/00Y/62)

(4) BARE PAST PARTICIPLE *DONE/BEEN/SEEN/GONE*:
  a. I never ***seen*** him. (SE/001/919)
  b. They ***been fixing*** the road. (SE/015/221)
  c. She ***gone*** to San Martin. (SE/005/114)
  d. Because what I had to do, I ***done*** it when I could. (SE/011/1144)

(5) *AIN'T* + VERB:
  a. He ***ain't wrote*** yet ... He ***ain't write*** yet. (SE/019/236-7)
  b. She ***ain't married*** none yet. (SE/005/160)
  c. I ***ain't got*** nothing to do. (SE/011/1143)
  d. I ***ain't*** never ***wore*** none. (ESR/00X/270)

This study considers the PERFECT in two corpora which represent an earlier stage of AAVE — Samaná English and the Ex-Slave Recordings. The Samaná English corpus comprises 21 interviews with native English-speaking descendants of American ex-slaves, who settled the remote peninsula of Samaná in the Dominican Republic in 1824 (Poplack and Sankoff 1987). The variety spoken by these informants is considered to derive from a variety of English spoken by African Americans in the early nineteenth century.[5] The Ex-Slave Recordings are a series of audio-recorded interviews with 11 former slaves born in the southern United States between 1844 and 1861 (Bailey et al. 1991). These corpora bear crucially and uniquely on the controversial origins and development issues in the current study of AAVE since they provide the necessary time-depth for assessing linguistic change (ca. 1800's) and the advantages of data drawn from naturally-occurring speech.

In PERFECT contexts, both Samaná English and the Ex-Slave Recordings exhibit the same forms attested in contemporary studies of AAVE, listed in (1)-(5) above. They also contain 'three verb clusters' with auxiliary *be* and *have*, as in (6), English preterite morphology, as in (8), and solitary verb stems, as in (7).

(6) THREE VERB CLUSTER WITH AUXILIARY *HAVE*:
   a. He told me that he ***had done pass*** through them English books. (SE/006/315-6)
   b. He ***had done been*** to Saint Thomas and place. (SE/001/647)

(7) THREE VERB CLUSTER WITH AUXILIARY *BE*:
   a. They ***ain't paid*** us yet and ***I'm done spent*** plenty money with the documents. (SE/006/155-6)
   b. ***I'm done been*** over there plenty but I don't like it. (SE/005/312-3)

---

[5] For detailed background and justification for this contention, see Poplack and Sankoff 1987; Poplack and Tagliamonte 1989; Tagliamonte 1991; Tagliamonte and Poplack 1988.

(8) PRETERITE MORPHOLOGY (SUPPLETION AND INFLECTION):
   a. They all ***died*** out already. (SE/013/80)
   b. But I don't know what ***took*** her now. (SE/015/245)

(9) UNMARKED VERBS:
   a. I'***m got*** eighty - going on eighty five. I never ***put*** my foot to [an] obeah. I don't believe in that. (SE/002/1072-3)
   b. I never ***like*** the city. (SE/013/113)

In this article I perform a distributional analysis of the forms used for the PERFECT in these materials. The term PERFECT is employed to refer to the semantic functions which prescriptive English grammar has labelled 'present perfect' tense. The morphosyntactic constructions that occur within these contexts are referred to as surface 'forms'. I approach these data from two different perspectives. In the first I take the semantic functions of the English PERFECT as the starting point and examine the frequency and distribution of forms that occur there. In the second, I begin with the individual forms and investigate their co-occurrence patterns with a number of independent features of the linguistic environment.

In order to assess the grammatical function and/or functions of these forms, I draw comparisons with standard and vernacular varieties of English and English-based creoles while at the same time casting the analysis into the larger context of linguistic change. My results suggest that despite the multitude of different forms, their distribution in Samaná English and the Ex-Slave Recordings patterns in the same way as the English perfect. Co-occurrence patterns of the most frequent forms in past time reference contexts more generally provide additional support for this contention. Further, parallels not just in form, but also in function with earlier stages of the English language suggest that the non-standard variants can be interpreted as synchronic remnants. These findings corroborate the accumulating evidence from earlier independent analyses of Samaná English and the Ex-Slave Recordings.[6]

---

[6] Poplack and Sankoff 1987; Poplack and Tagliamonte 1989; 1991; 1994; Tagliamonte 1991; Tagliamonte and Poplack 1988; 1993.

2. Previous Analyses of the PERFECT in AAVE, Creoles and English

2.1. AAVE

The standard English PERFECT has generally been considered absent from the underlying system of contemporary AAVE (Fasold and Wolfram 1975: 65; Labov et al. 1968: 254; Loflin 1970; but cf. Rickford 1975 for an alternative perspective). Three types of evidence have been adduced in favour of this contention. First, the morphosyntactic construction ***have*** + past participle is said to be extremely rare. Second, verbs other than ***have*** appear in auxiliary position, as in (10).

(10) a. I ***was*** been in Detroit.
b. I ***didn't drink*** wine in a long time. (Labov et al. 1968: 254)

Third, past participles, e.g. *been, done*, sometimes occur without a preceding auxiliary, as in (11), and where they cannot be accounted for by deletion of an underlying *have*.

(11) a. He ***been know*** your name.
b. He ***been own*** one of those.

This means they cannot be interpreted as an English past participle in a present or past perfect construction (Labov 1972b: 53).

The explanation for these linguistic facts involves not only a rejection of the PERFECT as a category of AAVE grammar, but also a denial that the standard English distinction between the preterite and past participle exists. A single surface form with no auxiliary appears across the board whether it surfaces with the morphology of an English past participle, as in example (12a), preterite, as in (12b), or there is alternation between forms, as in (12c).

(12) a. He ***taken*** it.
b. He ***came*** vs.
c. He ***done*** it. vs. He ***did*** it.

Wolfram and Fasold (1974: 66) suggest that instead of a separate past participle in AAVE, there is a 'general past form' that encompasses a number of separate categorical distinctions in English, particularly the simple past and PERFECT.

But what underlying grammar produced these forms? Many researchers have suggested they derive from a creole system. Dillard (1972a) divides pre-verbal *done* into two separate categories, one with an auxiliary preceding, e.g. *He's done come*, and one with no auxiliary, e.g. *He Ø done come*, attributing this difference to the distinct sources of the respective forms — AUX + *done* being an English form and Ø + *done* a creole form. Fickett (1972) suggests that *been* and *done* represent specific time periods in the past, i.e. *done* for recent past, and *been* for remote past time. Although this particular function for *done* is not widely attested, the remote time interpretation for *been* is quite widespread (e.g. Dillard 1972a; 1972b; Stewart 1965; Wolfram and Fasold 1974).[7]

2.2. Creoles

In creoles, pre-verbal *done* and *been* are widely-cited as typical tense/aspect features. While *done* is considered a perfective or completive marker (e.g. Alleyne 1980), *been* is considered a past and/or anterior marker, often with a remote interpretation (e.g. Agheyisi 1971; Faraclas 1987). The English *have* + past participle does not appear at all, pointing to a polar distinction between English and creole grammars (Bickerton 1975: 128). Unfortunately, the literature on this subject is entirely qualitative making form/function inferences about these forms difficult to assess. The only empirical investigation, Winford's analysis of the PERFECT in Trinidadian Creole, corroborates Bickerton's claim with its dramatic split between *have* usage with middle class speakers and verb stem forms with working class speakers (Winford 1993).

---

[7] Rickford 1975 specifies that the remote time interpretation is only applicable to the stressed version of *been* in AAVE.

2.3. English

But what exactly is the nature of the *English* PERFECT system? Much of the research claiming that AAVE has a creole-like grammatical system has based its conclusions on comparisons of AAVE features with standard (prescriptive) contemporary English usage rather than with vernacular varieties of English to which Africans must have had closer historical connections (Montgomery and Bailey 1986: 13), or with related present-day white vernaculars (Butters 1989: 194; Rickford 1990; Vaughn-Cooke 1987: 68) to which it might be more appropriately compared. Research on present-day varieties of vernacular American (Christian et al. 1988; Feagin 1979) and British English (Ihalainen 1976) as well as other regional varieties, e.g. Tristan da Cunha (Scur 1974) and Newfoundland, Canada (Noseworthy 1972)) have confirmed that many morphosyntactic forms used in PERFECT contexts in AAVE also appear in a wide geographic range of English dialects, many of which are entirely beyond the realm of creole influence. Thus, for example, there is no independent validation of Winford's (1993) claims that the patterns of surface forms used for PERFECT functions in Trinidadian Creole differ from an English one.

In what follows I describe the inventory of surface forms that have been attested in the literature on different varieties of English and review the hypotheses (where they exist) which have been put forward to explain them. We will see that the surface *forms* found in contexts of PERFECT reference are virtually the same across descriptions of AAVE, creoles and other varieties of English.

2.3.1. Have Deletion

The most frequently-cited non-standard form in PERFECT contexts is an English past participle which surfaces with no preceding auxiliary, as in (13) and in example (4) above.

(13) a. He ***been*** there. (SE/001/189)
b. Don't do that. I never ***done*** it. (ESR/008/25)

This form is attested in the United States (e.g. Atwood 1953; Christian et al. 1988; Fries 1940; Krapp 1925; Marckwardt 1958; Mencken 1971; Menner 1926; Vanneck 1955), Canada (Orkin 1971), Australia (Turner

1966), England (Wakelin 1977), Ireland (Visser 1970) and Tristan da Cunha (Scur 1974). The most popular explanation for this form is the *have*-deletion hypothesis which assumes that the forms with, and without, *have* fulfil the same function and thus can be attributed to the removal of an underlying *have* (e.g. Barber 1964; Wright 1905). But this does not explain its appearance in contexts in which the distinction between preterite and past participle appears to be neutralized (e.g. Menner 1926).

2.3.2. Generalized Past Marker

Thus, a second hypothesis for the bare past participles is that they represent the development of a new semantic category. They were originally based on the PERFECT but contexts in which the auxiliary syncopated, i.e. *I('ve) seen, I('ve) done*, led to complete elision. This auxiliary-less form was then adopted in vernacular varieties, reanalyzed as a preterite and extended to all the functions of the past tense (Menner 1926: 238; Vanneck 1955), so that *I seen him* has come to have exactly the same meaning as *I saw him*. (Mencken 1971: 520). This explanation for the bare past participle parallels the 'general past form' posited for AAVE (Wolfram and Fasold 1974: 66).

2.3.3. Loss of the PERFECT Tense

Some researchers have suggested this is the first stage in a process which will lead to the eventual loss of the PERFECT category in the grammar. This conclusion is said not to be surprising in light of the fact that the position of the PERFECT in the history of many languages is rather unstable, having been lost and reintroduced at various times (Scur 1974: 22; Vanneck 1955). For example, in French the gradual relaxation of the degree of recentness or current relevance required for use of the PERFECT enabled its form to supplant the simple PAST while losing its original meaning.[8]

---

[8] French, High German and Russian have all lost the distinction between preterit and perfect and the same phenomenon is characteristic of some other Germanic languages, Swedish and some Slavic languages as well (Scur & Svavolya 1975).

2.3.4. Lexical Restriction

Given these descriptions of *have* deletion one would think that bare past participles are frequent and productive forms. In fact, a bare past participle is a rare item in English since the only contexts where one can be unambiguously identified are with strong verbs. Weak verbs, which have no distinction between preterite and past participle morphology, would appear as preterites in the event of *have* deletion, making them indistinguishable from the (simple) past tense. Even within this limited range of contexts however, bare past participles rarely occur. An empirical study of variant forms of the PERFECT in the English of Tristan da Cunha, a small island in the South Atlantic, (Scur 1974) revealed only five — the verbs *see*, *be*, and *do* and sometimes *come* and *get*, as in (14) below.

(14) a. I ***been*** to South Africa.
b. We never ***seen*** a tractor around.
c. They ***done*** away with it.
d. We ***got*** plenty of them.
e. They just ***come***.

The same lexical restriction appears to be true of different varieties of English in England. Cheshire (1982) reports that working class teenagers in Reading used *done* categorically in the preterite, as in (15), while Hughes and Trudgill (1979) report variable occurrence of *seen*, (16a), and *done*, (16b), as preterites.

(15) She ***done*** it, didn't she, Tracy? (Cheshire 1982)

(16) a. You never ***seen*** them, you know. (Hughes and Trudgill 1979: 68)
b. I ***done*** another couple of years there, then they closed up. (Hughes and Trudgill 1979: 79)

2.3.5. *Been* and *Done*

Two frequently-cited bare past participle forms, *been* and *done*, require special mention because they appear in contexts which are not always directly translatable into standard English via *have* deletion (see Section

2.1). Despite this fact, these forms are attested in ***vernacular*** (white) English in a wide range of geographic locations in the United States and Canada (Feagin 1979; Noseworthy 1972; Williams 1975). ***Done*** has been referred to as an adverb (Feagin 1979) or a quasi-modal (Christian et al. 1988) and is generally considered 'completive/emphatic'. ***Been*** is generally attributed with meanings equivalent to the standard English PERFECT although in Newfoundland, e.g. (17a-b), Noseworthy suggests that it has a connotation of remoteness, indicating that the state of affairs took place 'farther back in the past than any action denoted by ... ***have*** + past participle' (1972: 21-2). Note the similarity to attested creole patterns (see Section 2.2). In Alabama, as in (18a-c), the meaning corresponds to 'begun in the past long ago and continued up to the present', or simply 'once, long ago', as in (18a-b).

(17) a. I ain't ***been done*** it.
b. I ***been cut*** more wood than you. (Noseworthy 1972: 22)

(18) a. I ***been knowin'*** your grandaddy for forty years.
b. Well, I chewed tobacco some, and then I started smokin' — started smokin' cigarettes. Course I — I ***been quit*** about 15 years since I smoked. (Feagin 1979: 255-6)

2.3.6. Three Verb Clusters

Although relatively obscure, three verb clusters, of the form AUX + ***done*** + verb, are also attested in vernacular (white) English in the United States (McDavid and McDavid 1960). Christian et al. (1988: 43) describe an uninflected form, i.e. ***done***, which occurs before an inflected verb optionally preceded by an inflected auxiliary in Ozark and Appalachian English, as in (19a-d). The same structure surfaces in Alabama English, as in (19e) (Feagin 1979).

Ozark English:
(19) a. I think they ***done took*** it.
b. Them old half gentle ones has all done ***disappeared***. (Christian et al. 1988: 33)

Appalachian English:

c. She asked us if we turned in the assignment; we said we done turned it in.

d. ... because the one that was in there ***had done rotted.*** (Christian et al. 1988: 33)

Annistan, Alabama English:

e. You buy you a little milk and bread and you'***ve done spent*** your five dollars! (Feagin 1979: 122)

2.3.7. Auxiliary *Be* vs. *Have*

Use of ***be*** as an auxiliary in PERFECT contexts instead of ***have*** is attested in contemporary varieties in England, Scotland, Ireland (Curme 1977; Edwards and Weltens 1985) and in the southern United States (Feagin 1979), as in (20).

(20) a. Some of the unions ***is done gone*** too far.

b. It was so quiet I thought everybody ***was done gone*** to bed. (Feagin 1979: 127)

2.3.8. Present Perfect vs. Simple Past Tense

Clearly, there is robust variability amongst PERFECT forms in vernacular English. In addition, although the meaning of past and perfect tenses in English is distinguished in many cases, researchers widely acknowledge that, even in the standard language, as in (21), (Quirk et al. 1985: 191; Wright 1905: 298) there are many contexts in which either one may be used (Frank 1972: 81; Leech 1971: 43).

(21) a. Now, where ***did*** I ***put*** my glasses?

b. Now, where ***have*** I ***put*** my glasses? (Leech 1987: 43)

This is also typical of Samaná English and the Ex-Slave Recordings, as in (22), where the past and perfect forms can occur within the same context, in the same discourse, by the same speaker, as in (23).

(22) a. God ***left*** me here for some purpose. (SE/002/390)
b. They ***didn't send*** it to me yet. (SE/022/390)
c. They all ***died*** out already. (SE/014/80)

(23) But the wind and the rain ***has wash*** them away. The rain ***wash*** them away. (SE/020/262-4)

In fact, in earlier varieties of English, interchangeability between these two categories was quite common and, in fact, far *more* variable than in the contemporary system. So, while many researchers have used distributional asymmetries with standard English functions to argue for an alternative grammar for surface forms used in PERFECT contexts, diachronic evidence may suggest another explanation. I now turn to the historical record.

3. Historical Development of the Perfect in English

In Old English, there were only two tenses: past and non-past. While the non-past served for durative and non-durative present and future reference, the past covered not only what is represented by the simple past of today, but also durative past tense (e.g. past progressive), as well as the PERFECT and past perfect tenses of the contemporary system (Strang 1970: 311). In other words, there were no overt forms to distinguish between habitual and progressive aspect and between PERFECT and NON-PERFECT meaning (Traugott 1972: 90-1). This can be seen in example (24) below, where habitual activity has no representative auxiliary (24a), and (24b) in which the simple past tense inflection marks a function that today would be overtly marked with the auxiliary and tense inflection combination of the perfect.

(24) a. 7 se cyning 7 ða ricostan men *drincaþ* myran meolc
'and that king and those richest men *drink* mare's milk'. (Traugott 1972: 89)

b. ðe cyðan hate ðæt me *com* swiðe oft on gemynd hwelce wiotan iu wæron giond Angelcynn
to-you tell command that to-me came very often to mind what wise-men before were throughout England
'Let it be known to you that it has very often come to my mind what wise-men there were formerly throughout England.' (Traugott 1972: 91)

Clearly, simple past and the perfect tenses were not differentiated. Moreover, it was often the case that the preterite forms marked a function that today would be overtly marked with the auxiliary and tense inflection combination of the perfect (Brunner 1963: 86; Traugott 1972: 90-91). In fact Visser (1970) claims that the simple past and present perfect are interchangeable in most contexts, including those where either one or the other alone would be required in contemporary usage.

During the change from Old to Middle English this two-tense (past vs. non-past) inflectional verb system underwent substantial elaboration (Strang 1970: 98), putting in motion a four-century long changeover from a highly inflectional or (synthetic) tense system to a periphrastic (analytic) one (Traugott 1972: 110).

3.1. Elaboration of the Verb Phrase

One of the most important changes to take place in the English time reference system was the development of separate elements within the verb phrase, in addition to the suffixal inflection on the main verb, to mark tense and/or aspect distinctions in addition to the original, and far more general, PAST tense.

3.1.1. *Have/had*

Perhaps the most prominent expansion of the tense system was the development of the present and past perfect tenses from the stative main verbs *have* and *be* as in (25) below.

(25) I *have* the letter written (i.e. in a written state).

Because the simple past tense gradually shifted in emphasis to explicitly PAST time there was a need for a new verbal structure that could function to represent a close relationship between PAST and PRESENT time. Since a written state implies a previous action, the structure *have written* gradually acquired verbal force, serving as a verbal form pointing to the past and bringing it into relation with the present (Curme 1977: 358).

During the initial phase of this development *have* and *be* competed as auxiliaries for the new category, as in (26); however, *have/had* gradually generalized to more and more verbs and eventually prevailed over *be* (Curme 1977: 359).

(26) a. He took his wyf to kepe whan he ***is gon*** *vs*. and also to ***han gon*** to solitaire exil

b. the yonge sonne ***hath*** in the Ram his halfe cours ***yronne*** *vs*. as rody and bright as dooth the yonge sonne that in the Ram ***is*** foure degrees up ***ronne***

(these examples from Chaucer cited from Brunner 1963: 87)

3.1.2. Three Verb Cluster

During the Middle English period a 'three-verb structure' developed, e.g. *He had done speak* (cf. Visser 1969: 338ff). While the origins of this form are obscure, it clearly represented a completed past time reference action, as in (27a). Inflection on the past participle was apparently variable as the form of the main verb originally surfaced as an infinitive, e.g. *speak*, but was gradually replaced by the past participle, e.g. *I had done spoke*, probably by analogy to forms such as *I done it* (Visser 1970: 2210). Similarly, as Traugott (1972: 146, n.18) points out, the past participle inflection *-ed* on weak verbs is not required. Hence forms such as *has done invent* and *has done invented* were synonymous, as in (27b).

(27) a. Also he seyde ... he ***hadde don sherchyd*** att Clunye.
'Also he said ... he had done searched at Cluny.'
(He had finished searching) (Traugott 1972: 146)

b. And many other false abusion The Paip (=Pope) ***hes done invent.***
(Traugott 1972: 146)

Between Middle and Modern English the form with *done* became stigmatized as nonstandard. It did not survive past the fifteenth century in Southern England (Williams 1975: 273); however, in the Northern dialectal regions it remained common.

3.1.3 Summary

The obvious similarities between the 'creole' forms reported in the literature on AAVE and these Early Modern English analogues has not gone without notice (e.g. Christian et al. 1988; D'Eloia 1973; Herndobler and Sledd 1976; Schneider 1993; Traugott 1972). The same forms as well as standard English *have* + past participle are also attested in written representations of earlier varieties of AAVE (Schneider 1989).

Comparisons based on similarities between surface forms alone however, do not provide unambiguous evidence for semantic function or genetic relationship. It is by now well-known that linguistic items from one language may pattern entirely according to another's rules (e.g. Bickerton 1975; Mufwene 1983a; Rickford 1977; Singler 1990; Tagliamonte et al. in press; Winford 1985). Other forms may represent two systems simultaneously (e.g. while verb stems in creoles have very similar interpretations to the English simple past tense, the same past tense can also be used interchangeably with the present perfect in many PERFECT contexts). Unfortunately, very few conditioning factors, in particular linguistic ones, which might help to illuminate these facts have ever been mentioned, nor, in the rare cases that such factors have been considered, have they ever been identified. Thus, there is no basis from which to differentiate between verbal patterns that are inherent to the English language and those which could possibly be due to hypercorrection, incomplete acquisition or even an alternate system. The case of the PERFECT in English and creole grammars is a particularly difficult site for disentangling these issues because it is a semantic domain in which there is a complete lack of isomorphism between morphological distinctions (i.e. form) and semantic distinctions (i.e. function).

4. Circumscribing the Variable Context

The conceptual space of PERFECT comprises both a semantic aspect (i.e. current relevance) and a semantic tense (i.e. indefinite past). Thus, the form *have* + past participle is related to more than one semantic function. On the other hand, what is often not recognized in the literature, is that these semantic functions can be represented by more than one *form* as well.

In addition to the parallels between *overt* English and creole PERFECT markers, both grammars can be expected to admit morpho-syntactially unmarked verbs for the same semantic functions. Because English (at least) has widespread phonological deletion in (weak) past time reference, verb stems are possible variants of the simple past. By extension, this means that in PERFECT contexts as well, at least three surface forms might occur: *have* + past participle, preterite and, to some extent, verb stems. In creoles, on the other hand, where the PERFECT is said not to exist, neither as the form *have* + verb, nor as a category in the grammar (Bickerton 1975: 129), we might expect either many verb stems in PERFECT contexts, as found by Winford (1993), and/or creole forms, such as *done* and *been*. Thus, as found in previous studies of the tense/aspect system, (Poplack and Tagliamonte 1989; 1991; 1994; Tagliamonte and Poplack 1993; Tagliamonte et al. in press) the mere existence of a form is not sufficient to identify the underlying grammatical mechanism that produced it.

Take, for example, the *been* + verb construction: If this surface form was produced by an English grammar it would be explained as one in which the auxiliary *have* has been deleted and would be construed as a variant of the PERFECT. While this form does correspond in some instances to the English present perfect as in, e.g. *John been workin' here all day today*, there are often cases where it corresponds to the past or past perfect tense as well, suggesting that it cannot be solely equated with the PERFECT and hence cannot be attributed to an English-like grammar (Bickerton 1975; 1979; Dillard 1972a; Mufwene 1983b; Stewart 1970). Instead, it may represent a creole remote past or anterior marker. Similar arguments can be made for the *done* + verb construction. It corresponds sometimes to English present perfect and sometimes to past perfect tense depending on the context (Mufwene 1988: 258) and for these reasons it may reflect an underlying creole function, such as completive, unrelated to the standard English system. However, differentiation between patterns that are inherent to the English language and those which derive from an alternate grammatical system can only be observed through analysis of the frequency and distribution of forms across all the contexts in which they might have occurred and in relationship to *all* other forms and functions within the past time reference system more generally.

5. Results

In order to evaluate these possibilities, the analyses reported here approach these data from two different perspectives — surface form and semantic function. First, every verb which referred to (realis) past time was extracted and coded for its morphosyntactic characteristics. Then, using prescriptive English grammar as point of comparison, each surface form was categorized according to the semantic tense/aspect function(s) for which it was used. This allows for a calculation of form/function correspondences in the data. Finally, the co-occurrence patterns of each surface form were examined according to a number of independent linguistic features from the literature on this subject, e.g. time adverbs, conjunctions, and remoteness.

5.1. Distributional Analysis by Semantic Function

Table 1 depicts the overall distribution of surface forms across all past time reference contexts. Observe that both Samaná English and the Ex-Slave Recordings have the same range of variants and, with no substantial exceptions, the same hierarchy.[9] As illustrated earlier, in (1)-(9), both contain surface forms consistent with the literature on the PERFECT in creoles as well as vernacular and historical varieties of English. *Have* + past participle and bare past participles occur in both

---

[9] The small differences in hierarchy amongst the rarer variants are undoubtedly due to their extremely low frequency overall.

Table 1
Overall distribution of surface forms found in past time reference contexts in ***Samaná English*** and the ***Ex-slave Recordings.***

Surface Form	Samaná English		Ex-Slave Recordings	
	%	N	%	N
Preterite	62	4861	58	1162
Verb stem	17	1311	16	331
Habitual, progressive etc.[10]	15	1152	18	371
was/got passive	2	150	2	47
had + past participle	1.5	120	1	15
have + past participle[11]	1	86	1	18
Past participle	.7	58	1	28
Verbal *-s*	.6	46	1	25
be + verb	.5	39	.04	1
ain't + verb	.5	36	.3	5
done + verb	.1	10	.4	7
had + *done* + verb	.07	6	.2	3
be + *done* + verb	.05	4	—	0
TOTAL N		7879		2013

corpora with the same frequency. *Done* + verb, as well as three verb clusters with auxiliary *be* or *have* also occur. But none of these surface forms exceed 1% of the data, not even the English PERFECT marker *have*. Can the striking infrequency of *have* forms be used as evidence that PERFECT is not a full-fledged category in these data? And is there any evidence that any of these fulfill creole-like, rather than English-like functions?

---

[10] This category consists of habitual forms such as *used to*, *would* + verb and variants of the progressive, e.g. *was going*, which are not the focus of this investigation.

[11] This includes *have/has/'s* as well as a following verb form which could include unmarked weak verbs and strong verbs with preterit morphology, in addition to standard English past participles.

These questions can only be answered by taking into account the distribution of forms by semantic function. For example, even though a surface form may be infrequent, this may be entirely due to the fact that the ***meaning*** which it embodies was also quite rare. Each past time reference verbal construction was coded according to all tense/aspect categories which could have been used in the same context: (i) the context required the present perfect, as in (28), and (ii) the context permitted either the present perfect or the simple past, as in (29) and (22) above, (iii) the context required the simple past, as in (30), (iv) the context required the past perfect, as in (31), and (v) the context permitted either the past perfect or simple past, as in (32). The remainder under the heading 'Other' consist of contexts permitting habitual and progressive forms which are not the focus of this study (cf. Tagliamonte and Poplack in progress).

(28) PRESENT PERFECT TENSE REQUIRED:
  a. But today we ***calmed*** off and everything ***got*** calm. (SE/002/942)
  b. I came in last Friday and I ***ain't been*** nowhere. (SE/002/1339-40)
  c. Now, those things ***fell*** out. (SE/016/173)

(29) PRESENT PERFECT OR PAST TENSE:
They ***didn't send*** it to me yet. (SE/001/1149)

(30) PAST TENSE REQUIRED:
This morning we ***went*** to the church in Clara. (SE/006/1549)

(31) PAST PERFECT TENSE REQUIRED:
Because they ***hadn't cut*** the road yet. (SE/002/708)

(32) PAST PERFECT OR PAST TENSE:
Well then, they killed the boy. After they ***killed*** the boy.... (SE/002/948)

Samaná English and the Ex-Slave Recordings represent an earlier variety of English spoken in the United States. If that variety developed directly from contact with contemporaneous English vernaculars, then it would not be unreasonable to expect that verbal constructions which have since disappeared from contemporary standard English might appear there. I hypothesize that if a specific set of surface forms was once possible in the semantic context for PERFECT, i.e. *have* or *be* auxiliary, three verb clusters, *done/been* + verb etc., then we should observe some proportion of each of these forms within those contexts. We should also expect restricted usage of some forms in environments which have become specialized to only one tense in contemporary standard English, a context which requires the present perfect for example. If, Samaná English and the Ex-slave Recordings are creole-like, on the other hand, then it would not be unreasonable to expect verb stems, *been* and/or *done* to appear in PERFECT contexts rather than *have*. Moreover, we should also expect the distribution of these forms to follow attested creole patterns, such as remoteness distinctions. Such correspondences will enable us to evaluate whether or not the distribution of morphological marking parallels what would be expected in a English or creole system.

Tables 2 and 3 (see over) depict the percent distribution of each surface form by semantic function. Note the infrequency, but highly partitioned distribution of the rarer PERFECT forms.[12] Bare past participles, preverbal *done*, auxiliary *be* and the three verb clusters are restricted to environments where the English present perfect tense can occur (or in the case of the three verb cluster with *had*, past perfect tense). The specifically creole form *been* + verb does not occur at all!

---

[12] Passives and verbal *-s* clearly pattern with the simple past tense. The latter are undoubtedly Historical Presents in the narrative complicating action section of narratives of personal experience. *Ain't* + verb is vanishingly rare and not specific to any context. See Howe 1994 for the absence of *ain't* in past, as opposed to present time reference contexts contra DeBose 1994.

Table 2
Percent distribution of surface forms by semantic function in *Samand English*.

SURFACE FORMS	PAST	PAST/ PAST PERFECT	PAST PERFECT	PAST/ PRESENT PERFECT	PRESENT PERFECT	OTHER	TOTAL N
	%	%	%	%	%	%	
Preterite	86	3	0.2	2	0.2	10	4861
Verb stem	83	2	—	4	0.4	10	1311
Habitual and progressive	18	0.002	—	0.8	—	81	1083
had + past participle	26	47	18	3	—	7	120
got passive	95	2	—	2	—	—	88
have + past participle	2	1	—	51	44	1	86
was passive	92	3	—	2	—	3	62
Past participle	19	9	—	45	26	2	58
3Vb cluster *had*	—	33	50	17	—	—	6
Verbal *-s*	41	—	—	—	4	54	46
be auxiliary	3	3	—	67	28	—	39
ain't	17	—	—	42	31	11	36
done + verb	20	—	—	60	20	—	10
3Vb cluster with *be*	—	—	—	50	50	—	4
TOTAL N	5728	221	33	277	96	1524	7879

Table 3
Percent distribution of surface forms by semantic function in the ***Ex-Slave Recordings***.

SURFACE FORMS	PAST %	PAST/ PAST PERFECT %	PAST PERFECT %	PAST/ PRESENT PERFECT %	PRESENT PERFECT %	OTHER %	TOTAL N
Preterite	72	1	0.17	0.86	0.09	26	1162
Verb stem	63	2	0.3	0.9	—	34	331
Habitual and progressive	14	—	—	0.28	—	86	360
had + past participle	—	53	47	—	—	—	15
have + past participle	—	—	—	39	61	—	18
was passive[13]	98	2	—	—	—	—	43
Past participle	25	14	—	11	46	4	28
3Vb cluster *had*	—	75	25	—	—	—	3
Verbal *-s*	76	8	—	—	—	16	25
be auxiliary	—	—	—	100	—	—	1
ain't	40	—	—	40	20	—	5
done + verb	—	14	14	29	43	—	7
3Vb cluster *be*	—	14	14	29	43	—	7
TOTAL N	1176	37	12	29	29	730	2013

---

[13] The *got* passive does not occur in these data.

Consider these patterns in the context of the history of the English language. The present perfect tense developed over a long period of time in which alternation of *have* and *be* as auxiliaries and even multiple auxiliaries such as *have + done* and *be + done* are amply attested. The sporadic, but localized occurrence of *exactly* the same forms here and in the very contexts where they would be expected to occur given this history is striking.

Historical grammars reveal that at least some aspects of the linguistic environment exerted an influence on the occurrence of some of these forms. The auxiliary *be* tended to be used with intransitives (Brunner 1963: 87) and where the participle clearly expressed the idea of a state or had an adjectival interpretation (Curme 1977: 359). Accordingly, we examine the distribution of auxiliary *be* according to the lexical aspect of the verb, illustrated in Table 4.

Table 4
Percent distribution of *be* vs. *have* auxiliary forms by lexical aspect in ***Samand English***.

SURFACE FORM	STATIVE		PUNCTUAL		TOTAL
	%	N	%	N	
be + verb	71	27	29	11	38
have + verb	54	46	46	40	86

Observe that verbs with a stative reading have a marked tendency to occur with auxiliary *be*. Moreover, 79% of these contexts were intransitive, as in (33). This patterning is identical to that suggested in the historical record.

(33) a. 'Cause them now, since the war ***is got*** civilized. (SE/002/746-7)

b. I'***m*** never ***been*** in prison half an hour. (SE/021/988)

Consider the bare past participles. The vast majority occur in contexts of present perfect tense, providing initial support for an underlying auxiliary. However, a non-negligible proportion (about 25%), occur in contexts for the simple past. Is this evidence for loss of the PERFECT via a past verb form generalizing across the verbal delimitation paradigm?

Further examination of these forms by lexical type, depicted in Table 5, reveals that bare past participles are restricted to only four verbs — *done*, *been*, *gone* and *seen*.

Table 5

Percent distribution of bare past participles by lexical type across semantic contexts in ***Samand English***.

SURFACE FORMS	PAST	PAST/ PAST PERFECT	PAST PERFECT	PAST/ PRESENT PERFECT	PRESENT PERFECT	OTHER	TOTAL N
	%	%	%	%	%	%	
been	—	—	—	60	40	—	25
done	43	17	—	35	—	4	24
gone	—	—	—	40	60	—	5
seen	25	25	—	—	50	—	4
TOTAL CONTEXTS	5728	221	33	277	96	1524	58

But it is actually only *done* and *seen* which occur in contexts of simple past, as in (34).

(34) a. They say they ***done*** as I ***done***. (SE/006/256)
b. The daughter came and she ***seen*** about her. (SE/003/443)

Moreover, the form and its function parallel present-day varieties of English (see Section 2.3.4). Thus, systematic encroachment of the bare past participle into the domain of simple past tense (see Section 2.3.3) is not borne out by these data.

In fact, present perfect contexts bear close to the entire inventory of *have* + verb forms in Samaná English, whereas this form is used only 1% of the time anywhere else. A similar pattern is found in the Ex-Slave Recordings. Preterite morphology, on the other hand, occurs very frequently, but only in the semantic contexts which require it in standard English. This leaves the bare stem form. Does its use reflect a creole grammar?

Clearly, its patterning is parallel to the *inflected* preterite form. Taking into consideration the fact that simple past tense is often rendered by the stem form due to phonological reduction processes in vernacular varieties of English (e.g. Guy 1980; Neu 1981) as well as in contemporary AAVE (e.g. Fasold 1972; Labov 1972b; Wolfram 1969), this parallelism of preterite and verb stem is entirely predictable. There is no association of the verb stem with PERFECT contexts as has been found in a creole system (see Section 2.2).

5.2. Summary

There are amazing parallels in the frequency and distribution of surface forms used for past time reference in Samaná English and the Ex-Slave Recordings. Those typical of contemporary English are the predominant forms in every one of the semantic contexts considered and their marking patterns are as would be expected in a English time reference system. While there are a number of *non-standard* forms, all of these have been previously attested both in the history of the English language or in dialects of contemporary English. Moreover, their functions, as can be determined here by the semantic contexts in which they occur, and by the other forms with which they are used, pattern according to what would be expected in an English grammar.

5.3. Distribution Analysis of Co-occurrence Patterns

I now turn to a distributional analysis of the most frequent forms[14] and examine their co-occurrence patterns across a number of independent features of the linguistic environment which are specifically related to

---

[14] The infrequency of the rarer surface forms do not permit comparable analysis.

PERFECT. I hypothesize that if a specific surface form is associated with a given feature in English (or creoles) and the same is found to be true in Samaná English and/or the Ex-Slave Recordings, then that will provide a point of comparison. If such parallelisms can be found across a number of features, I take this as evidence for similarity of the underlying grammatical mechanism regulating the distribution of forms in the data, and thus their grammars.

5.4. Temporal Distance

In creoles past time reference forms have been linked to relative distance from speech time. In English grammar differential location in time cannot be said to be relevant to any tense, except one — PERFECT — which occurs under conditions of recency and current relevance (Dahl 1984: 118). In order to determine the pertinence of temporal distance to the appearance of surface forms in Samaná and the Ex-Slave Recordings each verb was categorized according to the event time. For example, three distinct time periods are represented by the verbs in example (35): a remote time represented by the verbal structure *did buy*, in (35a), a less remote past time represented by the verbal structure *had went*, in (35b), and a comparatively recent past represented by two unmarked verbs, *come* and *stay*, in (35c-d).

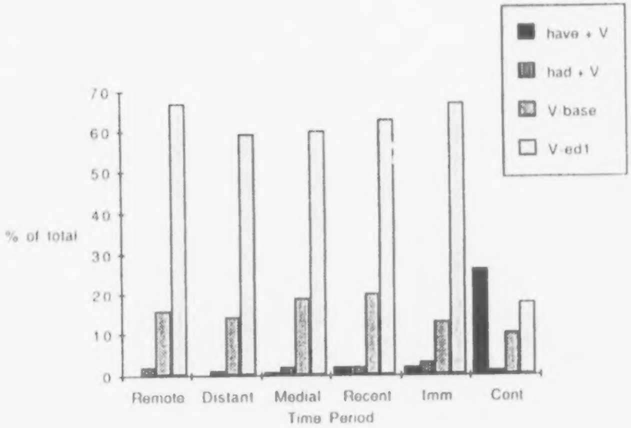
(35) a. But *in that time* we **did buy** sugar four cent the pound, you hear, four cent the pound, *time of Trujillo*.
b. And *from since that* look, the sugar **had went up** even to thirty cents, you hear.
c. And it **come** back *now* to twenty and eighteen.
d. And **stay** so, you hear. (002/890)

If the underlying system of these varieties is creole-like, we would expect to find a correlation between specific time periods and specific surface forms whereas if the system is English-like, the only area in which temporal distance will demonstrate an effect will be in immediate or continuing past contexts.

Figures 1 and 2[15] compare the distribution of surface forms across reference points at different temporal intervals in the past, i.e. remote, distant, medial, recent, immediate and continuing. These are given in terms of their percent occurrence out of the total number of all tense/aspect forms.

Figure 1

Distribution of preterite, verb stem, *have* + past participle, and *had* + past participle by time period — ***Samaná English*.**

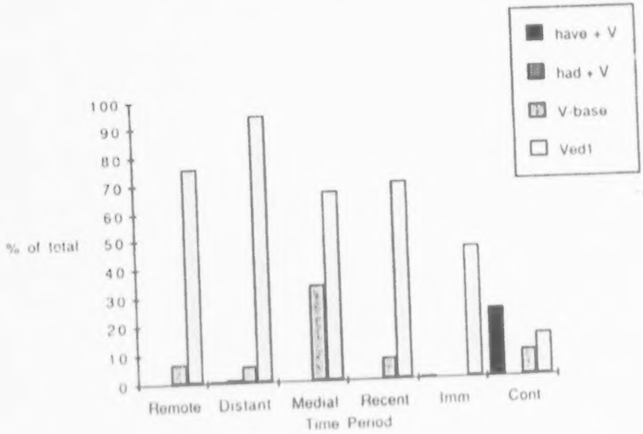


---

[15] Abbreviations used in the tables and figures can be interpreted as follows: 'V-ed1' refers to inflected or suppletive preterit forms. 'V-b' refers to a verb stem. 'V-s' refers to unambiguous present morphology, e.g. *don't*, *-s* . 'Hab' refers to habituals such as *used to* + verb and *would* + verb, among others. 'V-ing, 'Ø-ing' and 'is V-ing' refer to variants of the progressive.

Figure 2

Distribution of preterite, verb stem, *have* + past participle, and *had* + past participle by time period — ***Ex-Slave Recordings.***



Despite a skewed representation of temporal distance in the Ex-Slave Recordings,[16] all surface forms exhibit parallel occurrences across past time reference time periods. These distributional facts suggest that there is no remoteness distinction in the past time reference system of either of these varieties.

One temporal context is an exception, that of 'continuing past'. In both corpora it is composed of the same forms, *have* + past participle, preterite and verb stems, and in the same proportions. *Have* + past participle is almost non-existent in all other past reference times. The same pattern is evident in the Ex-Slave Recordings.

---

[16] In the Ex-slave Recordings, 94.2% of all verbs considered come from the same time period — that of the 'distant past'. This is the time period of the Ex-Slaves' youth and/or childhood from which most of their reminiscences take place. All other time periods combined make up only 122 tokens.

Recall that the function of the present perfect tense in English is to describe 'an alliance between past and present time' (Jespersen 1964). In these data, a form identical to that used in English for PERFECT distinguishes itself from other potential past time reference forms of sufficient frequency by the restriction of its occurrence to functions which have been identified throughout the prescriptive and historical literature on English as typical of PERFECT. Such correspondence between form and function can hardly be coincidental and I interpret this as another piece of evidence that the English present perfect is a viable tense/aspect category in Samaná English and the Ex-Slave Recordings.

5.5. Temporal Indicators

The interpretation of surface forms in creoles, particularly with regard to time reference, is said to be dependent on context. In English, on the other hand the difference between tense categories, especially between PERFECT and simple past tense, is marked by co-occurrence restrictions with specific adverbs (e.g. *lately, so far, already, yet, up to now*, etc.) and conjunctions (e.g. *before, after, since*, etc.) (cf. Huddleston 1984: 158-9; Jespersen 1964: 243; Leech 1971; Quirk and Greenbaum 1972: 44; Quirk et al. 1985).

5.5.1. Adverbs

In English grammar features which predict where the present perfect is preferred to the simple past are related to temporal specification (Visser 1970: 2192). In creole grammars on the other hand temporal adverbs provide contextual cues which help to disambiguate morpho-syntactically unmarked verb in addition to the information provided by the stative/non-stative distinction (Mufwene 1983a).

Accordingly, temporal indicators in the immediate (sentential) environment of each past-reference form were tabulated in order to determine what effect temporal adverbs have on surface morphological forms in the two corpora. Figure 3 shows the frequency of adverbial specification across surface forms.

Figure 3

Percent frequency of adverbs by surface forms in ***Samaná English*** and the ***Ex-Slave Recordings***.

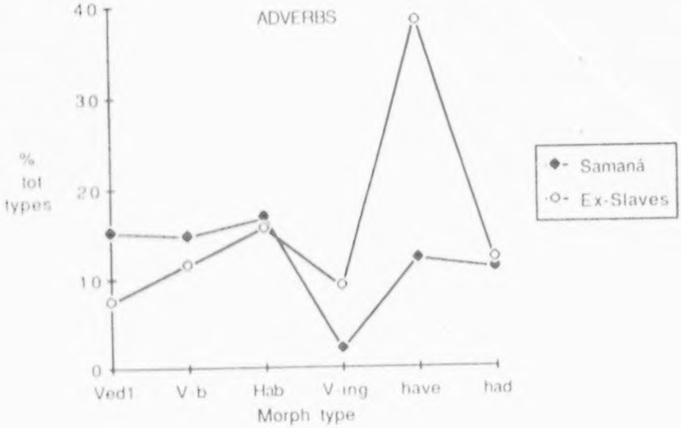


Figure 3 shows that the presence of a temporal adverb in the local clause structure has little effect on surface morphology in Samaná English. Marked and bare verbs behave almost identically. The high frequency of adverbs with *have* + verb in the Ex-Slave Recordings is due to the small number of contexts (N=18) in this category.

What happens when the adverbs are subdivided according to type? Prescriptive English grammar holds that some adverbs are linked to specific tense/aspect categories. For example, there is a restriction against the PERFECT with time-position adverbials referring to specific times, as in (36a). These adverbs, e.g. *yesterday, at that time, in 1901*, etc. force the occurrence of the simple past tense, as in (36b). Though not restricted to explicitly past time, time-frequency adverbials are said to occur with simple past tense forms which have a habitual semantic interpretation. Present relevance adverbs, on the other hand, i.e. those that refer to a period of time that stretches from a point in the past to speech time, (36b), are reserved for use with the present perfect (Visser 1970).

(36) a. *I *have seen* him last night.
b. I *have live* here twenty-one years. ... I came in the '61. (SE/019/82)

(37) *I BIN know you for a long time.

Tables 8 and 9 illustrate the distribution of surface forms by adverb type.

Table 8
Distribution of adverb types across surface forms in *Samaná English*.

	Preterite % (N)	Verb stem % (N)	Other % (N)	*have* % (N)	*had* % (N)	**Total** N
Time/ frequency	17 (17)	15 (15)	63 (62)	3 (3)	2 (2)	99
Time/ position	58 (143)	21 (53)	16 (40)	2 (5)	2 (6)	247
'then' (subsequence)	61 (212)	30 (105)	7 (26)	— (0)	1 (4)	347
Present reference	41 (18)	23 (10)	2 (1)	25 (11)	9 (4)	44
Continuous	27 (11)	41 (17)	41 (17)	15 (6)	— (0)	41
TOTAL						778

Table 9
Distribution of adverb types across surface forms in the ***Ex-Slave Recordings.***

	Preterite % (N)	Verb stem % (N)	Other % (N)	*have* % (N)	*had* % (N)	**Total** N
Time/ frequency	33 (16)	18 (9)	41 (20)	8 (4)	— (0)	49
Time/ position	69 (44)	13 (8)	11 (7)	11 (7)	2 (1)	64
'then' (subsequence)	32 (14)	39 (17)	30 (13)	— (0)	— (0)	44
Present reference	50 (1)	— (0)	— (0)	50 (1)	— (0)	2
Continuous	14 (45)	16 (5)	3 (1)	29 (9)	6 (2)	31
TOTAL						190

'Present reference' adverbs, illustrated in (38a-b), are distributed across all the surface forms but they are the *only* type of adverb that occurs with any degree of frequency in contexts marked by *have* in Samaná English. Of all adverbs that occur with *have* + verb (N=25), 44% are of this type.

(38) a. They knocked that out. Everything **now** *have* change. (SE/003/827-8)
b. I'm sorry some of them *haven't reach* **yet** that you'd see them. SE/(009/346)

Unfortunately the Ex-Slave Recordings contain only two of these so a similar comparison is impossible. The high percentage of other adverb types co-occurring with this morphological form in the Ex-Slave Recordings is due to a large number of continuous adverbials, as in (39a-c), which are also consistent with the English present perfect.

(39) a. I ***ain't had*** no clothes to buy since I been on the project and I***'ve been*** on it, I think, 'bout nine - 'bout **eight or nine years** I believe. (ESR/00Z/98)
b. Then he died. He been dead **forty some odd year**. (00Z/75)
c. We ***been*** slaves **all our lives**. (008/188)

On the other hand *been* never occurs with time position adverbs. Recall that in AAVE there is a restriction against the use of stressed *BIN* with exactly these adverbs. This means that the 'absolute restriction' against continuous adverbs in AAVE in contexts such as *for a long time*, as in ***I BIN know you for a long time*** (Rickford 1977) does not hold in these data. This can be clearly seen in (39b-c) above from the Ex-Slave Recordings and in (40a-b) below from Samaná English. In contrast, forms with *have* rarely occur with adverbs referring to specific time, e.g. *last night*.

(40) a. ... *been raining* a good bit **all these days pass**. (021/581)
b. I can't hardly tell you 'cause it *been* **so long**. (020/18)

Finally, time-position adverbs in Samaná English are restricted to preterite or verb stem forms — 58% with preterite and 21% with verb stems. The same is true of the Ex-Slave recordings where 69% of all these adverbs occur with the preterite and 13% occur with verb stems.

5.5.2. Conjunctions

Conjunctions with disambiguating temporal value (Chung and Timberlake 1985: 209) also have specific collocation restrictions in English. For example, *since* actually requires the use of the present perfect, e.g. *He has been finished since last March*. Others, such as *when*, imply coincidence. While forms such as *after* can be used with either simple past tense or past perfect (Quirk et al. 1972: 339).

Accordingly, each context in these data was tabulated for its occurrence with temporal conjunctions, as illustrated in (41).

(41) a. I've seed covers **since** ***I've been*** big enough. (ESR/00W/334)

b. Oh he was so mean, fractious that-a-way, **when** he got ***drinking***. (ESR/00W/470)

c. Well then **after** they ***had*** that war, well then all had to go home. (SEC/004/401)

Figures 4, 5 and 6 represent how the three main conjunction types, *since*, *when* and *after*, are distributed across surface forms in Samaná English, the only data set where there are a sufficient number of temporal conjunctions to view patterns of co-occurrence. In Figure 4 *since* occurs with *have* + past participle and *had* + past participle, although more frequently with *have* + past participle, the form which most closely approximates the English present perfect. In Figure 5, *when* exhibits a propensity to appear with present V-ing forms. *After*, illustrated in Figure 6, is said to occur either with the simple past or the past perfect. Predictably it is found with preterites, verb stems and *had* + past participle as well as with habituals (e.g. *used to, would* etc.).

Figure 4
Percent occurrence of *since* with each surface form in *Samaná English*.

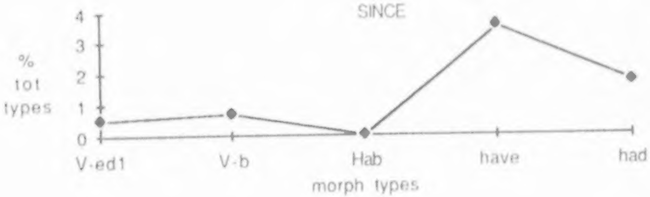


Figure 5
Percent occurrence of ***when*** with each surface forms in ***Samaná English***.

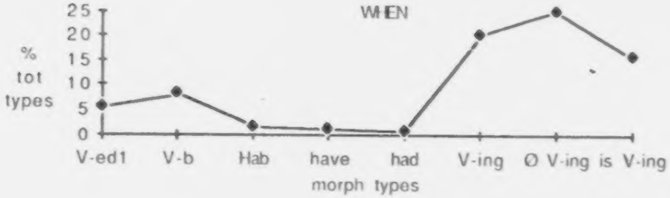
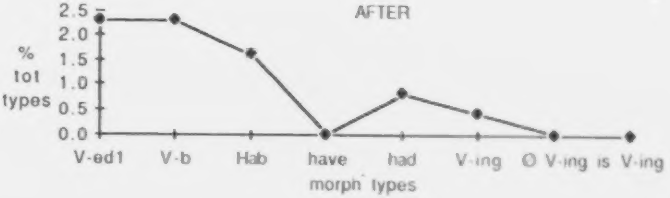


Figure 6
Percent occurrence of ***after*** with each surface form in ***Samaná English***.



5.6. Summary

Distributional analyses of co-occurrence patterns have revealed that the surface forms of past time reference are not differentiated by the relative 'remoteness' of past time except for that of 'continuing' past time. Here the context is restricted to *have* + past participle, preterite and verb stem. This behaviour parallels the English present perfect. Surface forms exhibit co-occurrence patterns similar to those of English. Time position adverbs co-occur with preterites and verb stems; time frequency adverbs co-occur with habitual and progressive forms. There are no functionally-motivated marked patterns as suggested for creoles in which morphosyntactially unmarked verbs would occur in contexts of temporal disambiguation. Present relevance adverbs are the only adverb type typical of the surface form *have* which is, once again, consistent with a PERFECT interpretation for the semantic function of this form.

The distribution of forms with conjunctions is also consistent with those suggested in English grammars, in which the surface form *have* + past participle patterns with *since*. Note also that the percent occurrence of preterite and verb stem is the same across all of the conjunctions considered here corroborating my earlier observations that these forms are variants of the same tense.

Although co-occurrence patterns such as these cannot be entirely conclusive in determining tense/aspect categories (cf. Comrie 1985), taken in conjunction with the partitioning of forms across semantic contexts, they provide corroborating evidence for interpreting these patterns as English-like, not creole-like, while at the same time confirming the parallelism between the two data sets more generally.

6. Discussion

This article has examined the PERFECT in Samaná English and the Ex-Slave Recordings through separate analyses of the distribution of forms by semantic function and co-occurrence patterns. Despite the overall rarity of this category in the general realm of past time, the most frequent forms used to mark it — *have* + past participle and bare past participles — are not at all marginal in contexts *licensed* for the present perfect in English. Co-occurrence patterns with temporal distance, adverbs and conjunctions also mirror those of the present perfect in standard English, while differing from those proposed for creoles. These findings suggest that the form *have* actually functions as a productive marker of PERFECT in these data. Bare past participles, with the exception of *seen* and *done*, are probably the result of *have* deletion since their occurrence is highly restricted to the same PERFECT contexts. Other surface forms attested in the literature were also found to mark PERFECT. Why should this be so?

I briefly outlined the development of the perfect in the history of English and found that it is perhaps the only tense/aspect category in English with such variability in forms. At its inception, auxiliary *had* and *be* were productive variants. In Middle English further elaboration of the verb phrase within the same domain of meaning led to the development of three-verb structures, *have/had done* + verb and *'m/is/are done* + verb. All these are attested in vernacular (white) English in the southern United States. As far as the bare past participle is concerned, forms such as *I seen/I done* have been traced to at least as early as the high tide of Irish immigration in the 1840's, the same time period represented by Samaná English and the Ex-Slave Recordings. In England they remain common in the West Midlands and the north and they also resemble Scottish forms. Thus, all the forms discussed here are found to persist in many contemporary varieties of English around the world where they are characterized as dialectal, non-standard, subject to style-shifting and the effects of education (e.g. Francis 1958). It is not surprising, given the extra-linguistic characteristics of the speakers in these corpora and the status of these varieties as linguistic enclaves, that members of an earlier English verb system persist, albeit marginally.

Is there any support here for the loss of the perfect? If *have* deletion is the first stage of this process, there is little evidence of a general process of change. While earlier studies have not provided actual figures for the frequency and distribution of *have* across lexical verbs, without evidence to the contrary we might assume that all verbs have an equal propensity to be used for PERFECT reference. But contexts in which *have* deletion occurs are restricted to infrequent realizations of *been, done* and *seen*. Infrequent preterite and verb stem forms in contexts of PERFECT cannot be taken to reflect either creole origins or ongoing change, since this usage is consistent with the historical record which documents extensive variation between preterite and present perfect tenses in earlier stages of English (see Section 3).

What of the forms that could not be subsumed under a *have*-deletion hypothesis? First, the creole-like structure *been* + verb did not even occur. Thus, of all surface forms found in these data, only those in

(3), namely *done* + verb, could not be interpreted as the deletion of a (standard) English past participle. Although these contexts are not structurally parallel to the contemporary standard English perfect, if we take the *three* verb cluster into account then these forms may simply be deletion of the same perfect auxiliary, but from a three place verb phrase, rather than the contemporary auxiliary + main verb structure. Thus, the *have* deletion hypothesis can be maintained.

The similarities between Samaná English, the Ex-Slave Recordings and other varieties of English and their lack of similarity with creoles can hardly be coincidental. Although English in the United States and the Caribbean could arguably have been influenced by creoles (but cf. Mufwene to appear-a; to appear-b for an alternative analysis) varieties such as those found in Newfoundland and Tristan da Cunha were not. Thus, the origins of these perfect forms and their functions must necessarily be traced to the original source in Britain. The rare PERFECT variants are remnants from an earlier stage in the development of the present and past perfect tenses in the history of the English language. Little, if anything, is known about the linguistic and extra-linguistic *conditioning* of variability in this area of the grammar. While the findings reported here now provide the basis for such analyses (Tagliamonte and Poplack 1995), it seems clear that the grammar of early Black English, insofar as it is instantiated by Samaná English and the Ex-Slave Recordings, was PERFECT just the way it was.

REFERENCES

Agheyisi, R. N. (1971) *West African Pidgin English: Simplification and simplicity*. Ph.D. Dissertation, Stanford University.

Alleyne, M. C. (1980) *Comparative Afro-American: An historical-comparative study of English-based Afro-American dialects of the New World*. Ann Arbor: Karoma.

Atwood, E. B. (1953) *A survey of verb forms in the Eastern United States*. Ann Arbor: University of Michigan Press.

Bailey, G., Maynor, N. and Cukor-Avila, P. (1991) *The emergence of Black English: Texts and commentary*. Amsterdam: John Benjamins.

Barber, C. (1964) *Linguistic change in present-day English*. Edinburgh/London: Oliver and Boyd

Baugh, J. (1980) A reexamination of the Black English copula. In W. Labov (ed.) *Locating language in time and space*. New York: Academic Press. 83-106.

Bickerton, D. (1975) *Dynamics of a creole system*. New York: Cambridge University Press.

Bickerton, D. (1979) The status of *bin* in the Atlantic creoles. In I. Hancock (ed.) *Readings in Creole Studies*. Ghent: E. Story-Scientia. 309-314.

Brunner, K. (1963) *An outline of Middle English grammar*. Oxford: Basil Blackwell.

Butters, R. (1989) *The death of Black English: Divergence and convergence in black and white Vernaculars*. 25. Frankfurt: Peter Lang.

Cheshire, J. (1982) *Variation in an English dialect: A sociolinguistic study*. Cambridge: Cambridge University Press.

Christian, D., Wolfram, W. and Dube, N. (1988) *Variation and change in geographically isolated communities: Appalachian English and Ozark English*. Vol. 74. Tuscaloosa, Alabama: American Dialect Society.

Chung, S. and Timberlake, A. (1985) Tense, aspect and mood. In T. Shopen (ed.) *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. 202-258.

Comrie, B. (1985) *Tense*. Cambridge: Cambridge University Press.

Curme, G. O. (1977) *A grammar of the English language*. Essex, Connecticut: Verbatim.

D'Eloia, S. G. (1973) Issues in the analysis of Nonstandard Negro English: A review of J. L. Dillard's Black English: Its history and usage in the United States. *Journal of English Linguistics*. 7.87-106.

Dahl, Ö. (1984) Temporal distance: Remoteness distinctions in tense-aspect systems. In B. Butterworth, B. Comrie and O. Dahl (eds.) *Temporal Distance: Remoteness Distinctions in Tense-aspect Systems*. Berlin: Mouton. 105-22.

DeBose, C. and Faraclas, N. (1994) An Africanist Approach to the Linguistic Study of Black English: Getting to the Roots of the Tense-Aspect-Modality and Copula Systems in Afro-American. In S. Mufwene (ed.) *Africanisms in African American Language Varieties*. Athens: University of Georgia Press. 364-87.

DeBose, C. E. (1994) A note on *ain't* vs. *didn't* negation in African American Vernacular. *Journal of pidgin and creole languages*. 9.127-30.

Dillard, J. L. (1972a) *Black English: Its history and usage in the United States*. New York: Random House.

Dillard, J. L. (1972b) On the beginnings of Black English in the new world. *Orbis*. 21.523-536.

Edwards, V. and Weltens, B. (1985) Focus on: England and Wales. In W. Viereck (ed.) *Varieties of English around the world* Amsterdam/Philadelphia: John Benjamins. 97-139.

Faraclas, N. G. (1987) Creolization and the tense-aspect-modality system of Nigerian Pidgin. *Journal of African Languages and Linguistics*. 9.45-59.

Fasold, R. (1971) Minding your Z's and D's: Distinguishing syntactic and phonological variable rules. In *Papers from the Seventh Regional Meeting, Chicago Linguistic Society*. Chicago: Department of Linguistics. University of Chicago. 360-367.

Fasold, R. (1972) *Tense marking in Black English: A linguistic and social analysis*. Washington, D.C.: Center for Applied Linguistics.

Fasold, R. and Wolfram, W. (1975) Some linguistic features of Negro dialect. In P. Stoller (ed.) *Black American English: Its Background and its Usage in the Schools and in the Literature*. New York: Dell Publishing Co., Inc. 49-83.

Feagin, C. (1979) *Variation and change in Alabama English: A sociolinguistic study of the White community*. Washington, D.C.: Georgetown University Press.

Fickett, J. G. (1972) Tense and aspect in Black English. *Journal of English Linguistics*. 6.17-19.

Francis, W. N. (1958) *The structure of American English*. New York: Ronald Press Co.

Frank, M. (1972) *Modern English: A practical reference guide*. Englewood Cliffs, N.J.: Prentice-Hall.

Fries, C. C. (1940) *American English grammar*. New York: Appleton, Century, Crofts.

Guy, G. (1980) Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.) *Locating language in time and space*. New York: Academic Press. 1-36.

Herndobler, R. and Sledd, A. (1976) Black English — Notes on the auxiliary. *American Speech*. 51.185-200.

Howe, D. (1994) *Patterns in the use of ain't in North Preston Vernacular English*. Manuscript. University of Ottawa.

Huddleston, R. (1984) *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Hughes, A. and Trudgill, P. (1979) *English accents and dialects: An introduction to social and regional varieties of British English*. London: Edward Arnold.

Ihalainen, O. (1976) Periphrastic 'do' in affirmative sentences in the dialect of East Somerset. *Neuphilologische Mitteilungen*. 77.608-622.

Jespersen, O. H. (1964) *Essentials of English grammar*. University of Alabama: University of Alabama Press.

Krapp, G. P. (1925) *The English language in America*. New York: The Century Company.

Labov, W. (1969) Contraction, deletion, and inherent variability of the English copula. *Language*. 45.715-762.

Labov, W. (1972a) *Language in the inner city*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1972b) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W., Cohen, P., Robins, C. and Lewis, J. (1968) *A study of the non-standard English of Negro and Puerto Rican Speakers in New York City*. Comparative Research Report. U.S. Regional Survey, Philadelphia.

Leech, G. N. (1971) *Meaning and the English verb*. London: Longman Group Ltd.

Loflin, M. (1970) On the structure of the verb in a dialect of American Negro English. *Linguistics*. 59.14-28.

Marckwardt, A. H. (1958) *American English*. New York: Oxford University Press.

McDavid, R. I. and McDavid, V. (1960) Grammatical differences in the north central states. *American Speech*. 35.5-19.

Mencken, H. L. (1971) *The American language*. New York: Alfred A Knopf.

Menner, R. J. (1926) Verbs of the vulgate in their historical relations. *American Speech*. 1.230-240.

Montgomery, M. B. and Bailey, G. (1986) Introduction. In M. B. Montgomery and G. Bailey (eds.) *Language Variety in the South: Perspectives in Black and White*. University: University of Alabama Press. 1-29.

Mufwene, S. S. M. (1983a) Observations on time reference in Jamaican and Guyanese creoles. *English Worldwide*. 4.199-229.

Mufwene, S. S. M. (1983b) Some observations on the verb in Black English Vernacular. *African and Afro-American Studies and Research Center Papers*. 5.1-46.

Mufwene, S. S. M. (1988) English pidgins: Form and function. *World Englishes*. 7.255-267.

Mufwene, S. S. M. (to appear-a) African-American English. In J. Algeo (ed.) *The Cambridge History of the English Language*.

Mufwene, S. S. M. (to appear-b) *The founder principle in creole genesis*. Diachronica.

Neu, H. (1981) Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In W. Labov (ed.) *Locating language in time and space*. New York: Academic Press. 37-54.

Noseworthy, R. G. (1972) Verb usage in Grand Bank. *Regional Language Studies Newfoundland*. 4.19-24.

Orkin, M. M. (1971) *Speaking Canadian English*. London: Routledge and Kegan Paul.

Pfaff, C. W. (1971) *Historical and structural aspects of sociolinguistic variation: The copula in Black English*. 37. Inglewood, California: Southwest Regional Laboratory Technical Report.

Poplack, S. and Sankoff, D. (1987) The Philadelphia story in the Spanish Caribbean. *American Speech*. 62.291-314.

Poplack, S. and Tagliamonte, S. (1989) There's no tense like the present: Verbal *-s* inflection in Early Black English. *Language Variation and Change*. 1.47-84.

Poplack, S. and Tagliamonte, S. (1991) African American English in the diaspora: The case of old-line Nova Scotians. *Language Variation and Change*. 3.301-339.

Poplack, S. and Tagliamonte, S. (1994) *-S* or nothing: Marking the plural in the African American diaspora. *American Speech*. 69.227-259.

Quirk, R. and Greenbaum, S. (1972) *A university grammar of English*. London: Longman.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1972) *A grammar of contemporary English*. New York: Harcourt Brace Javanovich, Inc.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. (1985) *A comprehensive grammar of the English language*. New York: Longman.

Rickford, J. (1975) Carrying the new wave into syntax: The case of Black English bin. In R. Fasold and R. Shuy (eds.) *Analyzing Variation in Language*. Washington, D.C.: Georgetown University Press.

Rickford, J. (1977) The question of prior creolization of Black English. In A. Valdman (ed.) *Pidgin and Creole Linguistics*. Bloomington: Indiana University Press. 190-221.

Rickford, J. (1990) *Grammatical variation and divergence in Vernacular Black English*. Paper presented at The 1989 ICHL workshop on internal and external factors in syntactic change.

Schneider, E. W. (1989) *American Earlier Black English, morphological and syntactic variables*. Tuscaloosa, Alabama: The University of Alabama Press.

Schneider, E. W. (1993) Africanisms in Afro-American English Grammar. In S. Mufwene (ed.) *Africanisms in Afro-American language varieties*. Athens: University of Georgia Press. 209-221.

Scur, G. S. (1974) On the typology of some peculiarities of the perfect in the English of Tristan da Cunha. *Orbis*. 23.392-396.

Singler, J. V. (ed). (1990) *Pidgin and Creole tense-mood-aspect systems*. Amsterdam/Philadelphia: John Benjamins.

Stewart, W. A. (1965) Urban Negro speech: Sociolinguistic factors affecting English teaching. In R. W. Shuy (ed.) *Social Dialects and Language Learning*. Champaign, Illinois: The National Council of Teachers of English. 10-18.

Stewart, W. A. (1970) Historical and structural bases for the recognition of Negro dialect. In J. E. Alatis (ed.) *Report of the 20th Annual Round Table Meeting on Linguistics and Language Studies*. Washington, D.C.: Georgetown University Press. 239-247.

Strang, B. M. H. (1970) *A history of English*. London: Methuen.

Tagliamonte, S. (1991) *A matter of time: Past temporal reference verbal structures in Samaná English and the Ex-slave Recordings*. Ph.D. Dissertation, University of Ottawa.

Tagliamonte, S. and Poplack, S. (1988) How Black English *past* got to the present: Evidence from Samaná. *Language in Society*. 17.513-533.

Tagliamonte, S. and Poplack, S. (1993) The zero-marked verb: Testing the creole hypothesis. *Journal of Pidgin and Creole Languages*. 8.171-206.

Tagliamonte, S. and Poplack, S. (1995) *The synchrony of obsolescence tracking the PERFECT in African Nova Scotian English*. American Dialect Society Meeting. Chicago, Illinois. December 1995.

Tagliamonte, S. and Poplack, S. (in progress) Habits in variation: The intersection of [+past]/[-punctual] in African American English in the Diaspora and its contact vernaculars.

Tagliamonte, S., Poplack, S. and Eze, E. (1996) Pluralization patterns in Nigerian Pidgin English. *Journal of Pidgin and Creole Languages*. 11(2).

Traugott, E. C. (1972) *A history of English syntax. A transformational approach to the history of English sentence structures*. New York: Holt, Rinehart and Winston.

Turner, G. W. (1966) *The English language in Australia and New Zealand*. London: Longmans.

Vanneck, G. (1955) The colloquial preterite in Modern American English. *Word*. 14.237-42.

Vaughn-Cooke, F. (1987) Are Black and White vernaculars diverging? *American Speech*. 62.12-32.

Visser, F. T. (1970) *An historical syntax of the English language*. Leiden: E. J. Brill.

Wakelin, M. F. (1977) *English dialects: An introduction*. London: Athlone.

Williams, J. M. (1975) *Origins of the English language: A social and linguistic history*. New York: Macmillan.

Winford, D. (1985) The concept of 'diglossia' in Caribbean creole situations. *Language in Society*. 14.345-356.

Winford, D. (1991) *Back to the past: The BEV/creole connection revisited*. Paper presented at Georgetown University.

Winford, D. (1993) Variability in the use of perfect *have* in Trinidadian English: A problem of categorial and semantic mismatch. *Language Variation and Change*. 5.141-188.

Wolfram, W. (1969) *A sociolinguistic description of Detroit Negro speech*. Washington, D.C.: Center for Applied Linguistics.

Wolfram, W. (1974) The relationship of White Southern speech to Vernacular Black English. *Language*. 50.498-527.

Wolfram, W. and Fasold, R. (1974) *The study of social dialects in American English*. Englewood Cliffs, New Jersey: Prentice-Hall.

Wright, J. (1905) *The English dialect grammar*. Oxford: Clarendon Press.

VOICE SOURCE CHARACTERISTICS OF MALE AND FEMALE SPEAKERS OF FRENCH.*

Rosalind A. M. Temple

University of York

1. Introduction

'Breathy Voice' is a phonation-type label used in phonology, in experimental phonetics and in speech pathology. 'Breathiness' is also a quality sometimes associated with females and with onsets and offsets of voiceless consonants. It is far from clear, however, what exactly are the acoustic characteristics of breathy voice, nor whether all the uses of the terms can properly be said to refer to the same phenomenon.

My purpose in the present article is to give a detailed account of part of an investigation into the realisation of the voicing contrast in plosive consonants produced by young French adults (Temple 1988a, b), which raised several questions which it was not possible to answer within the scope of that study, and to review the questions which arose at that time, in the light of subsequent literature.

2. Background to 1988 Study

2.1 The nature of 'breathiness'.

One physiological correlate of breathy voice quality is the vocal folds being held in the position for voiceless consonants, but the airflow rate is higher than normal and they vibrate loosely, 'so they appear to be simply flapping in the airstream' (Ladefoged 1982: 128), producing the breathy-voiced sound [h̤]. This occurs during the pronunciation of English intervocalic /h/, as in *ahead*. Another, more deliberate strategy is used in languages such as Gujarati, where there are phonemically contrastive breathy vowels, during which the vocal folds are held closely enough together at the front for voicing to occur, but apart at the back so that a large volume of air passes out through the glottis producing turbulence.

York Papers in Linguistics 17 (1996) 397-440
© Rosalind A. M. Temple

Bickley (1982) examined the vowels of Gujarati and !Xhóõ to determine acoustically and perceptually robust cues to the breathy-voice : modal-voice contrast. From the physiological description given in the previous paragraph one would expect an important cue to be the presence of high-amplitude inter-harmonic noise[1], and this is indeed found in the spectra of breathy sounds. However, following Ladefoged (1981) and other studies of Gujarati, Bickley wanted to investigate a cue at the other end of the spectrum, that of the relative amplitudes of the fundamental and the first harmonic above it[2]. She reanalysed Ladefoged's recordings of !Xhóõ and compared them with her own recordings of four native speakers of Gujarati. The measurements of the amplitude of the first two harmonics for the !Xhóõ speakers and one Gujarati speaker (*op. cit.:* 73-74) are reproduced as Tables 1 and 2 below. The figures show clearly that the fundamental (henceforth 'F0') is consistently higher in amplitude than the first harmonic above it (henceforth 'H2') in breathy vowels and not in clear vowels. To test the perceptual relevance of the cue, informal judgements were elicited from a native English speaker and a native Gujarati speaker, both trained in phonetics. The average amplitude differences for vowels judged to be in four categories of breathiness were as follows (the Gujarati speaker's judgements are given first): 'Very breathy' - 12.5dB, 10dB; 'Breathy' - 8.3dB, 11dB; 'Slightly breathy' - 6.7dB, 5.3dB; 'Not breathy' - 0dB, 0dB. Bickley synthesised /a/, /i/ and /u/ vowels with independent manipulation of the amplitude of the fundamental and the amount of aspiration noise, and the vowels were played to four Gujarati speakers . She found no correlation between the noise level and the degree of breathy percept, but the vowels with the highest amplitude F0 were consistently identified as breathy. Given the greater amount of noise passing through the glottis in breathy, as opposed to modal, phonation, it is surprising that the noise level did not have a greater effect on the breathy percept, but this may be because of problems with synthesis.

---

[1] Noise is the acoustic consequence of the turbulent airflow which would here be escaping between the parts of the vocal folds which are not fully adducted.

[2] The relative strength of the fundamental is known to increase as open quotient (the proportion of the vibratory cycle during which the vocal folds are open) increases. Increased open quotient is a known articulatory correlate of breathy voice quality.

	Difference (in dB)	
	Breathy	Clear
Speaker 1	13	0
Speaker 2	-4	-3
Speaker 3	2	-3
Speaker 4	5	-4
Speaker 5	5	-9
Speaker 6	4	-8
Speaker 7	11	0
Speaker 8	9	-2
Speaker 9	15	-2
Speaker 10	10	2

Table 1. Difference between amplitudes of first and second harmonics for breathy and clear vowels in !Xhóõ. After Bickley 1982: 73)

	Amplitudes in dB		
	first harmonic	second harmonic	difference
bar	44	42	2
maro	46	42	4
wali	47	43	4
bar	42	44	-2
maro	43	43	0
wali	38	44	-6

Table 2. Relative amplitudes (in dB) of first and second harmonics for breathy (top) and clear (bottom) vowels in Gujarati. (After Bickley 1982: 74)

"Breathiness" has also been much studied in a clinical context, sometimes being explicitly compared to the quality which is given the same label in other contexts. Hammarberg quotes a famous line of

Ladefoged's: '... what is a pathological voice in one language may be phonologically contrastive in another.' (Ladefoged 1983) and extends it to: 'What is evaluated as an abnormal voice quality in one language or dialect community may be a socially acceptable voice quality in another.' (Hammarberg, *op. cit.* 27) A particular spectral shape which is entirely attributable to physiological problems could thus be interpreted by speakers to convey a sociolinguistic message. Laver (1980) has exemplified how modes of phonation can be 'signals of emotional status' (Hammarberg, *op. cit.* 27) and Hammarberg's example is particularly pertinent to the present study, as we shall see in 2.2 below:

> 'For instance, breathiness is said to be a common female vocal attribute in many social communities, whereas creakiness often is a male characteristic.' (*ibid.* 27)

Hammarberg (1986) brings together a series of studies where pathological voices were judged by pathologists and phoniatrists against a series of voice quality parameters. The voices judged as breathy were all from patients with unilateral vocal-fold paralysis[3]. Acoustic analyses were made using long-term average spectra, and the typical long-term spectral characteristics of these voices were the high level of the fundamental, a low spectral level in the F1 region (400 to 600 Hz) and a high level of amplitude in the highest frequency band (5 to 10 kHz).

2.2 Female-male voice source differences

2.2.1 Acoustic evidence

The vocal folds of mature males are on average fifty per cent longer than those of females, and are thicker and greater in mass (Ohala, 1983). One natural result of this is that male fundamental frequency (F0) is lower than that of females[4]. As well as causing the perceived pitch of the male and female voices to be different, this difference in F0 means that the harmonics are more widely spaced and interact in a different way with vocal tract resonances[5]. Moreover, the shape of the female source waveform is more symmetrical than for males, and this is reflected in the amplitudes of equivalent harmonics, which decline more steeply in the case of the females. Monsen and Engebretson (1977) asked subjects to phonate into a long, reflectionless metal tube, which significantly reduced the resonances of the vocal tract and enabled them to analyse the glottal waveform. The waveform shape was found to be much more symmetrical for females than for males, with the opening and closing phases occupying almost equal proportions of the period. The male waveform had a characteristic 'hump' in the opening phase with the closing phase taking only twenty to forty per cent of the total period. These differences are reflected in the spectra with the slope in dB per octave between the harmonics being much steeper in the female glottal wave. The characteristics are not entirely surprising when one considers the physiology of the vocal folds: because of their greater mass, the males' vocal folds are drawn together faster than the females' by the Bernoulli effect, giving a sharper closure onset. Their larger size also results in the upper and lower parts being somewhat out of phase, which would create an effectively longer closure period. The waveform produced would thus be irregular in shape with enhanced harmonics above the fundamental. The female vocal folds, on the other hand, are drawn together less sharply, but with a smoother motion, and acting more as a single mass, which would produce a smoother, more sinusoidal waveform with the fundamental much stronger than the rest of the harmonics. Monsen and Engebretson's harmonic-by-harmonic comparison of glottal spectra in normal phonation (cf. Figure 1) reveals this difference in slope, but when the spectra are plotted un-normalised on the same frequency and amplitude axes, i.e. with the female signal about an octave higher in F0 than the male signal and with an overall intensity level -4 to -6 dB lower, the actual spectral envelopes are seen to be almost identical (cf. Fig 1b). There thus appears to be some sort of built-in normalisation factor for this particular spectral effect.

---

[3] Unilateral paralysis, and other deformations of the vocal folds, such as nodules, can impede complete closure during phonation, producing the same effect as in the normal speakers' production of breathy voice described above.

[4] Average values given by Fant (1956: 11, cited in Laver, 1983: 15) are 120 Hz for males and 220 Hz for females.

---

[5] The vocal tract resonances themselves are also different.

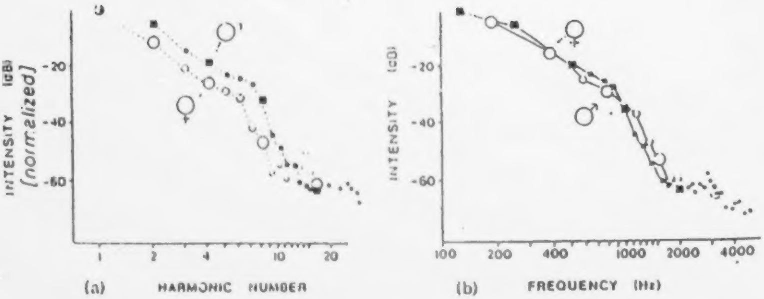


Figure 1. Average glottal spectra for male versus female normal voice phonation: (a) spectra normalised for both frequency and intensity; (b) non-normalised spectra. Male subjects, solid squares; female subjects, solid circles. (From Monsen and Engebretson 1977: 987)

It is interesting to note that when Bickley subjected steady-state vowels to inverse filtering to remove the effects of sound radiation and vocal tract filtering from the signal, her observations of the glottal waveforms produced in breathy and modal vowels corresponded closely to those observed by Monsen and Engebretson for female and male glottal waveforms respectively: 'The glottal waveforms of the clear vowels exhibited slower opening than closing phases, abrupt closure, and a closed phase that occupied approximately one third of the period of vibration.... The glottal waveforms for breathy vowels exhibited similar opening and closing phases, resulting in a more symmetrical shape. Closure was less abrupt and the closed interval was shorter.' (Bickley 1982: 76-77)

Other studies by those concerned with the synthesis of female-sounding speech confirm Monsen and Engebretson's findings concerning source differences. Klatt (1986) analysed the speech of a single female speaker with a 'pleasant voice quality'. He found considerable random breathiness noise above 2kHz over parts of many utterances and a variable degree of general tilt of the spectrum (i.e. over a larger frequency range than the F0-H2 measure) and of the strength of the fundamental. He attributes this variation to the presumed degree to which the larynx is spread or constricted.

2.2.2 Perceptual evidence

Barry (1986) reviewed some of the literature on male-female voice source differences and also concluded that they had much to do with physiology. His own study sought to make good synthetic copies of a male and female voice, to derive from these a set of tables that would reproduce the voice quality (using a rule-synthesis algorithm on the parallel-formant synthesiser developed by Holmes), and then to establish transformations which could be applied to one set of tables to derive the other. The acoustic features modified were F0, formant frequencies, spectral tilt and noise. In manipulating spectral tilt, Barry found that the best match was obtained by reducing the amplitude of the second formant (A2) by 6dB relative to the male A2, and of the third and fourth formants by 8dB. The male voice was generated without aperiodicity in the source signal (although there had been some present in the human subject) and this did not seem to make it sound unnatural. A 'good match' female voice included 25% noise. A discrimination test was carried out where listeners were played pairs of utterances and asked to select the one which sounded more like an adult female. The utterances most consistently judged as female were those where the formant frequencies *and* amplitudes and the spectral noise level of the original 'male' synthetic voice had been modified. It proved impossible to adjudicate between the relative importance of formant amplitude (and hence spectral tilt) adjustments and the degree of spectral noise. Thus, Barry's perceptual findings confirmed the importance of the production phenomena discussed in 2.2.1 above in the perception of a voice as "female".

2.2.3 Sociolinguistic claims

It would seem from the evidence just reviewed that the common claim that breathiness is a female attribute is predictable on the grounds that the physiology of female vocal folds gives rise to acoustic structures which are known to cue both a breathy and a "female" percept. However, the variability in degree of tilt found by Klatt suggests that although physiology (a constant for a given speaker) plays a significant role, voice source characteristics can be varied by manipulating the larynx constriction.[6] It is known from investigations into other acoustic phenomena that physiologically-predictable characteristics can be endowed with sociolinguistic significance by speakers and exaggerated or compensated for. For example, Mattingly (1966) tested the hypothesis that formant frequency differences between speakers of the same dialect were chiefly due to variations in the vocal tract size of the speakers, using data from Peterson and Barney's seminal 1952 study of vowels[7]. If the hypothesis were correct, Mattingly argued, there should be high correlation scores between the distributions of values for F1, F2 and F3 for the three classes of speaker (men, women and children). What he found in all but a few subsets of the data was that the correlation scores were in fact very low, and that the separation between male and female distributions of formants for some vowels was far sharper than could be explained by vocal tract size variation. He concluded:

> '... the difference between male and female formant values, though doubtless related to typical male and female vocal tract size, is probably a linguistic convention.'

Further evidence for the linguistic conventionalisation of cues to speaker identity which originated as physiological differences comes from work on children's speech before the development at puberty of physical vocal tract differences, since at the earlier stages there would be no physiological reason to account for sex-specific differences. Sachs (1975), for example, played children's' productions of /a/, /i/ and /u/ vowels to a panel of listeners, and asked them to identify the sex of the speaker. She obtained a statistically significant correct response rate of 66%, which suggests that the children (who were aged between 4 and 12) were beginning to produce sex-specific formant patterns despite the fact that the boys' and girls' vocal tracts would still be similar in size.

---

[6] If this were not possible, it would not be possible for female speakers of Gujarati and other languages, where breathy voice is used distinctively, to make the necessary distinctions. We shall return to this issue below.

[7] Peterson, G. E. & H. L. Barney (1952) Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24: 175-184

Vowel	/a/	/ɑ/	/ʌ/	/ɒ/
Females	8.4	6.4	6.2	3.3
Males	0.98	0.77	0.16	0.39
Difference (F-M)	7.42	5.63	6.04	2.91

Table 3. Average differences in amplitude in dB between the first and second harmonics in male and female speakers of Received Pronunciation. (After Henton and Bladon 1985: 224)

Henton and Bladon (1985) did not consider the physiological basis of source spectrum differences corresponding to breathiness, but they did examine the male-female differences as a sociolinguistically determined sex-specific marker. They followed Bickley (1982)[8] and measured the amplitude of F0 and H2 in the steady-state portions of open vowels produced by male and female RP and 'Modified Northern' speakers. Their results for the RP speakers are reproduced in Table 3. The male-female differences were significant according to a *t*-test ($p<0.01$) and the difference across all the vowels (mean of means) was 5.5dB. As Henton and Bladon point out (*op. cit.* 225), the differences 'would be sufficient to carry the perceptual contrast between breathy and modal vowels' for Bickley's Gujarati speakers; however, when their measurements are compared with the values of the synthetic vowels played in Bickley's perceptual experiment, it would appear that only /a/ would be considered as more than 'slightly breathy' by either of Bickley's phoneticians (compare Table 3 with the values given on p.2 above).

Interestingly, when Watson (1987) asked colleagues to listen to his child-subject's voices, they did not perceive them as breathy until the possibility was pointed out to them:

> 'It may be that we accept as normal in children what would be 'breathy' in adults, until we are specifically

---

[8] It should perhaps be noted that speaker sex was not specified by Bickley, but it is assumed, because of the consistency of her results, that her speakers were all male.

> called on as phoneticians to attend to phonation type.' (*ibid.* 21)

The comment could easily be applied *mutatis mutandis* to sex-specific differences in breathiness: might it not be the case that breathiness is a comparative measure to be assessed against the cultural norm for modal voice, and therefore cannot be measured in universal terms? Alternatively, it could be that although we are dealing with measures along the same acoustic continuum, it is unjustified to speak of what is being labelled as breathiness as being classifiable as exactly the same phenomenon in both the case of females (and children) and that of a linguistic phonation type. If there were no difference, Gujarati women would have great difficulty in producing phonologically breathy sounds which were sufficiently different from sounds phonated with their modal voice.

Henton and Bladon would presumably not consider these questions to be problematic, as they see the spectral tilt[9] characteristics as being produced deliberately by the British female speakers, rather than as being a result of physiology, and would presumably hypothesise that female modal voice would not have the same culturally determined properties in Gujarati. On the premise that breathy voice is used to convey intimacy in English (Laver 1980: 135) they suggest that the RP. speakers are trying to sound 'sexy' [*sic*]:

> 'At an ethological level, breathy voice may be seen as part of the courtship display ritual, as important as bodily adornment and gesture. A breathy woman can be regarded as using her paralinguistic tools to maximise the chances of her achieving her goals, linguistic or otherwise.' (*op. cit.* 226).

---

[9] Hitherto the term 'tilt' has been used in its generally accepted designation of the rate of decrease in amplitude across the whole source spectrum; I shall also be using the term in this article to refer to the difference in amplitude between F0 and H2. I make no claims as to the equivalence of these two measures, using the term in refer to this amplitude difference.

The claim that the female RP voice has the distinctive spectral characteristics described solely with the paralinguistic aim of aiding the speaker to attract a mate seems rather exaggerated, especially in the light of the other papers discussed above which hold that the female source spectrum would tend towards the 'breathiness' pattern anyway for physiological reasons. However, this does not rule out the role of other sociolinguistic forces which could cause female speakers to move nearer to or further away from the physiologically determined female 'norm', which is the implication of the findings of Mattingly, cited above. It should also be pointed out, of course, that males may well be modifying their voice quality for similar reasons.

2.3 Breathiness and the Voicing Contrast

As is well-known, French, like English, has a two-way 'voicing contrast' between cognate pairs of obstruents, but as far as plosives are concerned, the labels *'Voiced'* and *'Voiceless'*[10] correspond to different phonetic patterns of realisation in the two languages, most obviously in the timing of vocal-fold vibration relative to the release of the consonant when in absolute initial position. The *Voiced* plosives of French are canonically voiced throughout the closure and release period, usually with no break (though see Temple 1988a, b); *Voiceless* plosives have no prevoicing and little or no aspiration. English *Voiced* stops are phonetically voiceless unaspirated, while the *Voiceless* ones are voiceless and with longer aspiration following release. In addition to the timing of voicing relative to the release of the consonant, there are many other phonetic correlates to the voicing contrast in French and English plosives which are well-documented elsewhere and which it is not necessary to review here (see Temple 1988a for references). One correlate which has been less thoroughly documented, although it is taken to be a well-known fact about at least English plosives, is that *Voiceless* plosives tend to have breathy voice at vowel onset, due to the vocal folds' beginning to vibrate before being fully adducted for the vowel. Ní Chasaide and Gobl (1988) reported an analogous process during the pre-aspiration of plosives in Swedish. Laryngographic traces showed vibration of the vocal folds as they opened for the *Voiceless* plosive, and this was accompanied by an increase in spectral tilt. However, they also found that the onset of voicing in post-consonantal vowels was much less 'clean' than the breathy offset of the pre-consonantal vowel.

---

[10] The labels *Voiced* and *Voiceless*, in italics and with initial capital letters will be used throughout this paper to refer to phonological categories. The same words in non-italic script, and entirely in lower-case will be used to refer to the phonetic distinction between stops with prevoicing and those without. Henceforth no citation marks will be used.

The evidence reviewed thus far shows that F0-H2 differences have been found to correlate with perceived "breathiness" in languages where this quality plays a phonological role. The same measure has been found to differentiate male and female voice sources, and this is to some extent predictable from male-female physiological differences. Moreover, it has been suggested that variability in this measure could have a sociolinguistic value. Temple 1988a and 1988b thus attempted to draw together whether degree of breathiness, measured by the F0-H2 difference, was yet another marker of the voicing contrast in initial position, and whether there were differences between male and female French speakers, and if so, whether there was interaction between sex-specific and voicing-specific effects.

3. The 1988 Study

3.1 Methodology

Seven speakers were recorded in their study bedrooms at the École Normale Supérieure in Paris, and two at Oxford University Phonetics Laboratory (O.U.P.L.), reading lists of monosyllabic words with initial plosive consonants in isolation and in the frame, 'Jean avait dit ____'. The stimuli were presented individually on cards to minimise listing effects, and the first element of each list was discounted. The six plosive phonemes of French occurred several times before each of three vowels, /i/, /a/ and /u/. Only tokens with the vowel /a/ were measured for this part of the experiment because it is in here that the lower harmonics are least likely to be affected by the first formant, either in transition or in steady-state. The data were analysed using the Signal File Manager of O.U.P.L.'s New England Digital microcomputer (see Clark 1986 for details). Windows were positioned at the points indicated by the letters A to E and V in Figure 2, that is, in the relatively steady-state parts of

407

the pre-voicing and the vowel, over the release itself and over the pitch periods closest to the release. The two frames which fall into this latter category were at varying distances in milliseconds from the release: B covered the last three pitch periods of prevoicing for females and the last two for males, including cases of *Voiced* stops which were partially devoiced (i.e. where voicing ceased before release); and D covered the first three and first two periods after release in both *Voiced* and *Voiceless* stops, the latter having varying Voice Onset Times. The frame lengths of 20ms and 16ms for males and females respectively were chosen after experimenting to find settings which would give the best resolution of harmonics whilst maintaining comparable lengths in both time and number of periods. For each frame, frequency in Hz and amplitude in decibels (dB) of F0 and H2 were noted.[11]

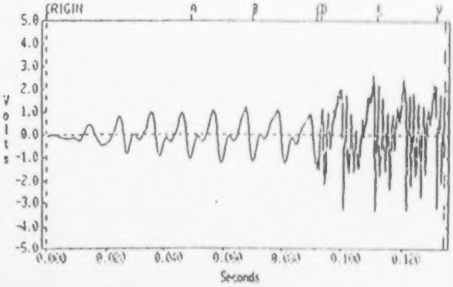


Figure 2. Positions of start of spectral windows for utterance "bac", by speaker PIG (male)

---

[11] For more details on the analytical procedure followed, see Temple 1988a: 57-70.

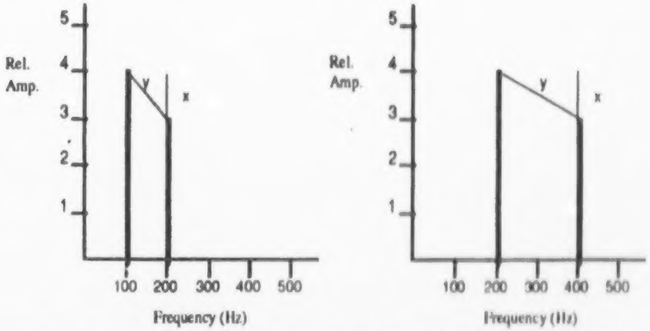


Figure 3. Schematic representation of the effects of fundamental frequency on the relations between harmonics in the spectrum.

A technical problem arises here in the question of how to compute what we have been referring to 'spectral tilt'. Both Bickley and Henton and Bladon calculated the straightforward difference between the amplitude measurements of the harmonics. Assuming that all Bickley's subjects were male, it is unimportant whether the measure is computed in this way or whether a true slope is calculated in amplitude loss per frequency unit (difference in dB 'over' difference in Hz). However, as soon as speakers with notably different F0 are to be compared, the choice of calculation method becomes important, since a higher F0 means a greater distance in Hz between F0 and H2, which would have a significant effect on the calculation of the slope. A schematic example is given in Figure 3 to illustrate this effect. The horizontal axis represents frequency in Hz, the vertical axis a hypothetical amplitude range. The solid vertical lines correspond to idealised harmonics for a male versus a female speaker. The difference in amplitude between F0 and H2 in both pseudo-spectra is 1. However, if the slopes are calculated in Amplitude/Frequency the results are 1/100 = 0.01 'A'/Hz for spectrum M, but 1/200 = 0.005 'A'/Hz for spectrum F. As well as having implications for comparisons across studies, this has implications for comparisons within a single study wherever speakers have significantly different fundamental frequencies. Indeed, spectra with a different amplitude difference could actually have the same slope gradient: if the difference in 'A' in spectrum M were 10, and in spectrum F 20, the gradients would be 10/100 = 0.1 'A'/Hz, and 20/200 = 0.1 'A'/Hz respectively. The question of which is the best way of measuring spectral 'tilt' is evidently potentially important and we shall return to it below. For the purposes of the experiment being described here it was decided to compute the measure both in terms of amplitude differences and in terms of dB/Hz slope.

Statistical analysis of the measurements was carried out using S.A.S.[12] Institute package implemented on the VAX mainframe computer at Oxford University Computing Service. The data were subjected to a 'General Linear Models' (G.L.M.) procedure, which allows Analysis of Variance to be carried out on 'unbalanced' models, because the numbers of tokens analysable for each speaker were not the same, principally because of the hazards of making recordings outside the recording booth.

3.2 Results and discussion.

3.2.1 Waveforms

No procedures were used to derive the source waveform from the vowel signal, but the waveforms during the closure period of prevoiced stops did appear consistently differently in male versus female subjects. Generally the waveform shapes in the speakers considered here seemed to be as predicted by Monsen and Engebretson, that is with a near-sinusoidal appearance for females, but with a 'hump' in the opening phase and a sharper closing phase for males (compare Figures 2 and 4).

---

[12] Statistical Analysis System.

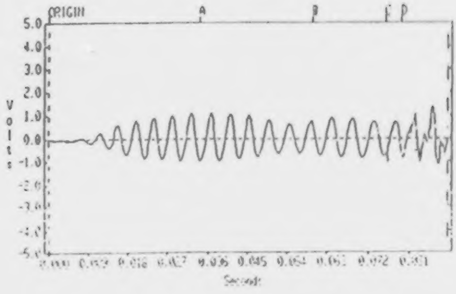


Figure 4. Waveform in prevoicing of "bac" by speaker ISR (female).

3.2.2 Relationship between F0 and H2: male versus female speakers

Position		A	B	D	E	V
Sex						
Males	dB/ Hz	-0.0378	-0.0813	-0.0262	-0.0093	-0.0183
Females		-0.0491$^{\$}$	-0.104$^{\$}$	-0.0492$^{\$}$	-0.0398$^{\$}$	-0.0396$^{\$}$
Males	dB	-5.026	-6.213	-3.758	-1.346	-0.404
Females		-15.853$^{\$}$	-18.330$^{\$}$	-10.642$^{\$}$	-8.920$^{\$}$	-9.504$^{\$}$

Table 4. Mean F0-H2 differences for frames positioned at A, B, D, E & V by male and female speakers expressed in terms of slope (dB/Hz) and amplitude (dB)

Mean values for the differences between F0 and H2 at the different positions in the word are given in Table 4 and Figure 5 in terms both of the dB/Hz slope and of amplitude comparisons in dB. A negative number indicates that the value for the fundamental is higher than for the second harmonic, and a positive number represents a lower value for F0. Another convention adopted has been to indicate the steeper gradient slope or greater amplitude difference in a particular two-way comparison with a superscript dollar sign ($^\$$). All the values in the table are higher for females than for males, as predicted from the evidence discussed hitherto, and the male-female contrast is high

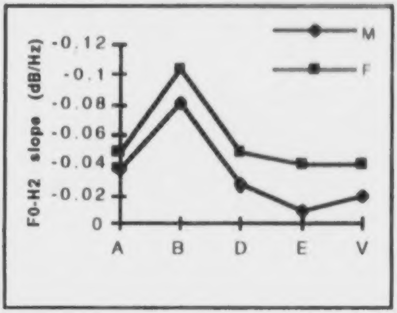


Figure 5a. Mean F0-H2 slope (dB/Hz) across positions of all tokens for males and females.

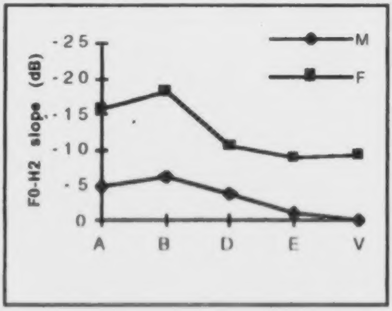


Figure 5b. Mean F0-H2 differences of amplitude across positions of all tokens for males and females.

significant according to a t-test ($p<0.001$) in all cases except V for the dB/Hz measure, which fails to reach significance even at the 5% level. It is clear from Figure 5a that the male and female trends in terms of slope stay firmly apart but follow much the same pattern with a sharp rise in steepness at B, that is as the release approaches or the prevoicing is about to cease. However, this effect is apparently reduced dramatically, particularly for females, in Figure 5b, where both curves are much smoother, showing only a slight rise in the dB difference at B. Also apparent in this Figure is the reflection of how the male-female difference at V 'becomes' statistically significant when calculated in terms of amplitude.

These findings are interesting for two particular reasons. Firstly, the only position where a significant difference was not found is the only one where measurements were taken in the other experiments reported, i.e. the relatively steady-state portion of the vowel. Secondly, they seem to confirm that changing the method of calculating the 'value' of the harmonic difference does have a significant effect on the apparent relationships between the sets of production data, which in turn suggests it could be relevant perceptually. Moreover, the measure which fails to reach statistical significance in this position is not the one used in the papers cited above, which begs the question 'how would those results look when calculated in these terms?'

3.2.3 Possible influence of consonant place of articulation

The steady-state part of the vowel was chosen by the other researchers referred to in order to avoid the possible effects of the F1 transition from the preceding or following consonant, which could enhance the amplitude of F0 or H2 and thereby distort the results. However, because the focus of this study was on the voicing contrast in consonants, these transition sections were precisely the parts of the signal in which we were interested. The only way to counteract the influence of formants would have been inverse filtering, which it was not possible to carry out at the time. Instead statistics were used to compare the effects of the different places of articulation of the consonants on the spectral values. Of course, the use of statistics cannot be seen as a replacement of inverse filtering by an equivalent measure, but we can hope that it would at least make us aware of any significant effect of components which would have been filtered out by that process. The slope values obtained for males and females are given in Table 5, and the amplitude-difference values in Table 6. Values are given for each position for each phoneme, and accompanying each value, an indication of those phonemes which are significantly different from the one in question, at the 5% level (t-test).

Position		**A**		**B**	
Consonant		*Mean*	*Diff From*	*Mean*	*Diff From*
/b/	*m*	-0.04348	g	-0.05464	
	f	$-0.09521^{\$}$	g	$-0.12022^{\$}$	g
	bth	-0.06430	g	-0.08147	g
/d/	*m*	-0.04975	g	-0.05778	
	f	$-0.09092^{\$}$	g	$-0.10448^{\$}$	
	bth	-0.06606	g	-0.07637	g
/g/	*m*	-0.01971	b d	-0.03248	
	f	$-0.06282^{\$}$	b d	-0.08860	b
	bth	-0.03781	b d	-0.05479	
/p/	*m*				
	f				
	bth				
/t/	*m*				
	f				
	bth				
/k/	*m*				
	f				
	bth				

Table 5(a). Mean slope differences (dB/Hz) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Positions A and B as in Figure 2.

Position		**D**		**E**	
Consonant		*Mean*	*Diff From*	*Mean*	*Diff From*
/b/	*m*	-0.01503	d	-0.01150	t
	f	-0.03099$^{\$}$	g k	-0.04675$^{\$}$	
	bth	-0.02134	d g k	-0.02545	
/d/	*m*	-0.04283	b g t	-0.01776	t
	f	-0.05104$^{\$}$		-0.03575$^{\$}$	
	bth	-0.04614	b t	-0.02502	
/g/	*m*	-0.02125	d	-0.01418	t
	f	-0.06695$^{\$}$	b	-0.03492$^{\$}$	
	bth	-0.03871	b	-0.02210	
/p/	*m*	-0.02941		-0.01638	t
	f	-0.04457$^{\$}$		-0.03897$^{\$}$	
	bth	-0.03611	b	-0.02637	
/t/	*m*	-0.01626	d	-0.00897	p b d g
	f	-0.04713$^{\$}$		-0.04004$^{\$}$	
	bth	-0.03056	d	-0.01375	
/k/	*m*	-0.03092		-0.00690	
	f	-0.05593$^{\$}$	b	-0.04233	
	bth	-0.04182	b	-0.02234	

Table 5(b). Mean slope differences (dB/Hz) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Positions D and E as in Figure 2.

Position		V	
Consonant		*Mean*	*Diff From*
/b/	*m*	+0.01744	p
	f	-0.04594$^{\$}$	
	bth	-0.00763	
/d/	*m*	-0.00461	
	f	-0.03259$^{\$}$	
	bth	-0.01591	
/g/	*m*	-0.05037$^{\$}$	
	f	-0.02796	
	bth	-0.04181	
/p/	*m*	-0.07412$^{\$}$	b
	f	-0.03895	
	bth	-0.05857	
/t/	*m*	-0.00814	
	f	-0.04275$^{\$}$	
	bth	-0.01544	
/k/	*m*	-0.01151	
	f	-0.04672	g
	bth	-0.02685	

Table 5(c). Mean slope differences (dB/Hz) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Position V as in Figure 2.

Position		A		B	
Consonant		*Mean*	*Diff From*	*Mean*	*Diff From*
/b/	*m*	-5.546	g	-7.044	
	f	-17.160$^{\$}$	g	-20.989$^{\$}$	g
	bth	-10.218	g	-12.749	g
/d/	*m*	-6.328	g	-7.268	g
	f	-16.435$^{\$}$	g	-18.754$^{\$}$	g
	bth	-10.331	g	-11.840	g
/g/	*m*	-2.762	b d	-4.040	d
	f	-11.682$^{\$}$	b d	-14.903$^{\$}$	b d
	bth	-6.506	b d	-8.359	b d
/p/	*m*				
	f				
	bth				
/t/	*m*				
	f				
	bth				
/k/					

Table 6(a). Mean amplitude differences (dB) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Positions A and B as in Figure 2.

Position		D		E	
Consonant		*Mean*	*Diff From*	*Mean*	*Diff From*
/b/	*m*	-2.362	d	-1.444	
	f	-7.236$^{\$}$	k	-9.158$^{\$}$	
	bth	-4.290	ptkd	-4.496	
/d/	*m*	-5.048	b	-2.466	t
	f	-10.343$^{\$}$		-7.309$^{\$}$	
	bth	-7.185	b	-4.421	
/g/	*m*	-2.896		-1.084	
	f	-11.085$^{\$}$		-7.255$^{\$}$	
	bth	-6.025		-3.442	
/p/	*m*	-4.586		-1.703	
	f	-10.339$^{\$}$		-9.467$^{\$}$	
	bth	-7.131	b	-5.137	
/t/	*m*	-2.846		-0.220	d
	f	-11.377$^{\$}$		-9.674$^{\$}$	
	bth	-6.799	b	-4.601	
/k/		-4.682		-1.168	
		-12.750$^{\$}$	b	-10.077$^{\$}$	
		-8.197	b	-5.050	

Table 6(b). Mean amplitude differences (dB) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Positions D and E as in Figure 2.

Position		V	
Consonant		*Mean*	*Diff From*
/b/	*m*	-0.093	
	f	$-10.317^{\$}$	
	bth	-4.137	
/d/	*m*	-0.797	
	f	$-7.573^{\$}$	
	bth	-3.532	k
/g/	*m*	-0.680	
	f	$-6.797^{\$}$	k
	bth	-3.017	k
/p/	*m*	+0.579	
	f	$-9.561^{\$}$	
	bth	-3.906	
/t/	*m*	-0.117	
	f	$-10.426^{\$}$	
	bth	-4.769	
/k/		-1.593	
		$-11.607^{\$}$	
		-5.955	

Table 6(c). Mean amplitude differences (dB) across place of articulation for the different sexes with indications of pair-wise contrasts significant at the 5% level (t- test). Position V as in Figure 2.

measure, but we can hope that it would at least make us aware of any significant effect of components which would have been filtered out by that process. The slope values obtained for males and females are given in Table 5, and the amplitude-difference values in Table 6. Values are given for each position for each phoneme, and accompanying each value, an indication of those phonemes which are significantly different from the one in question, at the 5% level (t-test).

Again, the dB/Hz slopes for females are consistently steeper than the males' slopes across all positions except at V for /g/ and /p/. The picture becomes more interesting when these values are compared with the dB values. For /p/, the male H2 is seen to be higher than F0. For /g/, both the measures show F0 generally higher than H2, but whereas the dB difference is greater for females than for males, with the other measure the result is the opposite. An extension of the hypothetical example above shows that this is mathematically unsurprising: with differences in 'A' of 10 in spectrum M, and of 20 in spectrum F, we saw that the gradients would be the same; however, a reduction of just one 'A' unit would give an apparently steeper slope for spectrum M, even though the amplitude difference would still be greater in spectrum F: 10/100 = 0.1 'A'/Hz; 19/200 = 0.095 'A'/Hz. Moreover, bringing the amplitude difference in spectrum F down to, say, 13 would still leave it greater than the difference for M, but in the slope would be 0.06 'A'/Hz, only just over half as steep as the male counterpart.

There are further differences between the two tables in terms of which pair-wise contrasts between phonemes show a significant difference. To take the values for the prevoicing first, although the 'Diff From' columns for measurements at position A are identical, there are discrepancies in the same column for position B, where, for example, /d/ enters into no significant contrasts for the dB/Hz measure, but contrasts with /g/ for all groups of speakers for the dB measure. With or without these discrepancies, these pair-wise contrasts also indicate that a caveat needs to be added to our suggestion above that the waveform of the prevoicing was the closest we were likely to get to the glottal source waveform. They show (not surprisingly) that the supralaryngeal characteristics of the consonants do affect the pre-voicing F0-H2 tilt. There are still large differences between males and females, but it could be argued that since place of articulation obviously does have an effect on the slope, the differences in the lower spectral components could be accounted for by supra-glottal differences, rather than differences generated by the vocal folds themselves. In view of the findings of the literature reviewed earlier, it is improbable that the male-female spectral differences found can be entirely ascribed to supra-glottal effects, but there was no possibility of testing the extent of those effects within the framework of this study.

In the post-release positions, the numbers of pairs of phonemes with significant differences between them decreases in both tables from D through E to V, but again different pair-wise contrasts were found to be significant in the different tables. It is clear too that the formant transitions do have an effect on the slope, and one is again forced to question whether the highly significant male-female differences found at D and E (as opposed to the failure to attain significance at V in the dB/Hz measure) were not at least enhanced by supraglottal resonance differences between the males and females. The effect of F1 would be reduced by the time it had passed through the frequency band where it would affect H2, hence the reduced inter-phoneme differences through E to V. If H2 is being enhanced, that would reduce the difference between it and F0, thus masking the characteristics of the 'breathy' spectrum. That there still is at least some male-female difference at V is encouraging for our original hypothesis that there is an effect independent of formant differences. However, this should be confirmed by examining the possible influence of the different F1 values of the vowels themselves. Actual measurements of the formant frequencies were not carried out, but a statistical analysis of possible vowel effects was done.

3.2.4 Possible effect of following vowel

Henton and Bladon (*op. cit.*) restricted their study to the English vowels /a/, /ɑ/, /ʌ/ and /ɒ/ in order to try and minimise the interference of F1 (which is relatively high in these vowels) with F0 or H2. The results comparing vowel-contexts for the present data in dB/Hz are given in Table 7 and Figure 6. Unfortunately the full set of statistics for the dB measure is not available, so in the light of the differences noted in the previous paragraph, the following comments, which are based on the dB/Hz values, should be taken with a note of caution.

Position		**A**		**B**	
Vowel		*Mean*	*Diff From*	*Mean*	*Diff From*
/i/	*m*	-0.04733	–	-0.04956	–
	f	-0.19102$^{\$}$	–	-0.10871$^{\$}$	–
	bth	-0.06517	e	-0.07338	e
/e/	*m*	-0.00096	–	-0.00485	–
	f	-0.06823$^{\$}$	–	-0.07630$^{\$}$	–
	bth	-0.02787	i u	-0.03036	i a u
/a/	*m*	-0.04061	–	-0.04887	–
	f	-0.07567$^{\$}$	–	-0.09836$^{\$}$	–
	bth	-0.05492		-0.06985	e
/u/	*m*	-0.03751	–	-0.05564	–
	f	-0.08800$^{\$}$	–	-0.11258$^{\$}$	–
	bth	-0.05738	e	-0.07758	e

Table 7(a). Mean slope values (dB/Hz) showing effects of different following vowels at positions A, and B across the sexes and indications of pair-wise contrasts significant at the 5% level (t-test figures for both groups only).

Position		**D**		**E**	
Vowel		*Mean*	*Diff From*	*Mean*	*Diff From*
/i/	*m*	-0.03399	–	-0.00966	–
	f	-0.08181$	–	-0.06914$	–
	bth	-0.05418	e a	-0.03477	e a u
/e/	*m*	+0.00739	–	+0.01403	–
	f	-0.07696$	–	-0.00605$	–
	bth	-0.02424	i u	+0.00650	i u
/a/	*m*	-0.01902$	–	-0.01444$	–
	f	+0.00132	–	-0.00339	–
	bth	-0.01028	i u	-0.00969	i u
/u/	*m*	-0.03485	–	-0.00576	–
	f	-0.07782$	–	-0.05574$	–
	bth	-0.05290	e a	-0.02675	i e a

Table 7(b). Mean slope values (dB/Hz) showing effects of different following vowels at positions D, and E across the sexes and indications of pair-wise contrasts significant at the 5% level (t-test figures for both groups only).

Position		V	
Vowel		*Mean*	*Diff From*
/i/	*m*	-0.03901	–
	f	-0.06499$^{\$}$	–
	bth	-0.04997	a
/e/	*m*	+0.00408$^{\$}$	–
	f	+0.02486	–
	bth	+0.01187	
/a/	*m*	-0.00409	–
	f	-0.01133$^{\$}$	–
	bth	-0.00766	
/u/	*m*	-0.02060	–
	f	-0.05256	–
	bth	-0.03402	a

Table 7(c). Mean slope values (dB/Hz) showing effects of different following vowels at position V across the sexes and indications of pair-wise contrasts significant at the 5% level (t-test figures for both groups only).

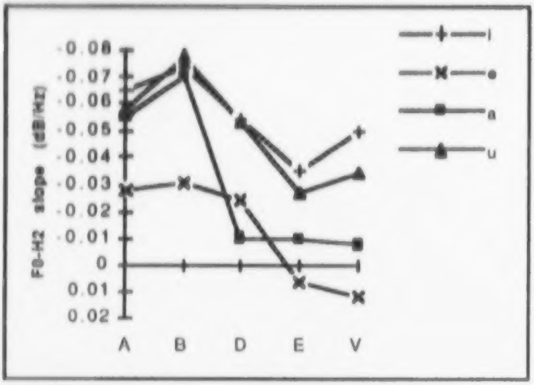


Figure 6. Slope differences as a function of following vowel. All speakers.

In Figure 6 the patterns for the four vowels when all speakers are taken together have a somewhat similar trajectory. Apart from /e/, there is a striking degree of similarity before the release, suggesting relatively little coarticulatory effect on this part of the spectrum in prevoicing. The atypical pattern for /e/ can be explained by the lack of tokens following either /b/ or /d/. There are large post-release differences and an inspection of the values for males and females separately (cf Table 7) shows that there is a complex effect, which is not surprising when one considers the complex sex-specific differences found in the acoustic structure of vowels. The female slope is again generally steeper. However, in /a/, where following previous studies we had expected to see the hypothesis confirmed most firmly, the male-female position is reversed after the release through D and E, and the only mean value for females to be a positive value (indicating H2 higher than F0) is at D for /a/ (although the male-female difference fails to reach significance at either D or E). At V there is a return to the more common pattern of females having the steeper mean slope, although this difference fails to reach significance by a long way (p>0.05). Clearly more detailed analysis of the interaction of slope and formant frequency is needed.

3.2.5. The Voicing contrast

It was suggested above that the F0-H2 difference may be found to vary following voiced versus voiceless consonants as an indicator of increased breathiness in the voiceless case. Values for the Voiced versus Voiceless classes as wholes are given in Table 8. None of the differences in slope between Voiced and Voiceless reaches significance. The greatest differences tend to occur in the vocalic portion, which is again where we should least expect to find them. The cross-phoneme comparisons shown in Tables 5 and 6 above revealed hardly any significant differences between cognate pairs, so these values are not surprising and no positive conclusions can be drawn from then concerning the discrimination of phonological classes.

	Position	D	E	V
Sex	**Voicing**			
m	*Voiced*	-0.02731$^{\$}$	-0.01466$^{\$}$	-0.01206
	Vless	-0.02509	-0.00415	-0.04039$^{\$}$
f	*Voiced*	-0.04945$^{\$}$	-0.03898	-0.03542
	Vless	-0.03579	-0.02040$^{\$}$	-0.03263$^{\$}$

Table 8. Mean values for F0-H2 slope (in dB/Hz) across Voicing categories for males and females at post-release positions.

If, as suggested above, this is not an effect manipulated by speakers but one due more to the physical effects of the gradual adduction of the vocal folds, we should expect the de-voiced tokens to follow the pattern of the *Voiceless* ones. Means were therefore computed across phonetic voicing type and are presented in Figures 7 to 9 and Table 9. Two graphs are given for the data for the male speakers and for the data for all speakers considered together because of the drastic effect of the mean V value for the O-PREV tokens. The categories represented are fully-voiced tokens (FVOICED); *Voiceless* tokens (PHON VLESS); *Voiced* tokens where prevoicing ceased at some time at or before release (DEVOICED); *Voiced* tokens with no actual prevoicing (0 PREV).

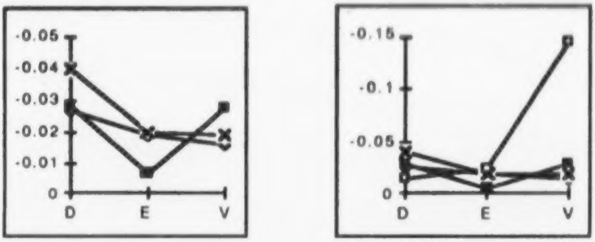


Figure 7. Slope values across positions for voicing type. Male speakers. Including (b), and not including (a), 0-prevoiced Voiced *tokens.*

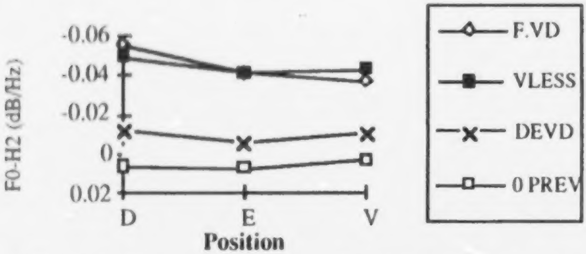


Figure 8. Slope values across all positions for voicing type. Female speakers.

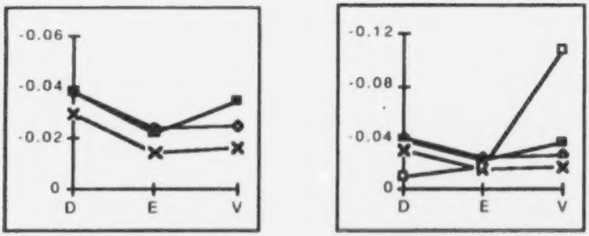


Figure 9. Slope values across positions for voicing type. Male speakers. Including (b), and not including (a), 0-prevoiced Voiced *tokens.*

Position		D	E	V
Sex	Voicing type			
	voiced	-0.02697	-0.01186	-0.01625
m	*Voiceless*	-0.02829	-0.00615	-0.02821
	devoiced	-0.03975	-0.01986	-0.01898
	Vd -- no prev	-0.01354	-0.02455	-0.14358
	voiced	-0.05509	-0.04092	-0.03706
f	*Voiceless*	-0.04896	-0.04039	-0.04275
	devoiced	-0.01121	-0.00546	-0.01030
	Vd -- no prev	+0.00628	+0.00687	-0.00290
	voiced	-0.03826	-0.02353	-0.02461
both	*Voiceless*	-0.03755	-0.02150	-0.03473
	devoiced	-0.02965	-0.01478	-0.01592
	Vd -- no prev	-0.00858	-0.01670	-0.10695

Table 9. Mean values for F0-H2 slope (in dB/Hz) across voicing categories for males and females.

When the effect of the male 0-PREV tokens is disregarded, the patterns for the different voicing types across the spectral window positions are very similar. There are no significant differences between types for males or for the group as a whole, but for females the FVOICED and the VLESS are significantly different from the DEVOICED and 0-PREV types, as reflected in Figure 9. With regard to the voicing contrast, therefore, there seems to be no phonetic or phonological grouping for which this measure of breathiness is a robust acoustic correlate.

4. Studies published since 1988

A good deal of work has been published since 1988 on the nature of voice source characteristics. We shall restrict ourselves here to a description of just a small number of important studies.

The most substantial single study is that of Klatt and Klatt (1990) on the analysis, synthesis and perception of voice quality variation. Klatt and Klatt analysed recordings of ten female and six male speakers uttering two 'real' sentences and reiterant imitations of those sentences using [ʔɑ] and [hɑ] syllables and measured the relative strength of the first harmonic, the presence of noise in the F3 region and above, and the presence of extra poles and zeros in the vowel spectrum, mid-way through the vowel. They found an average male-female difference of about 5.7dB in F0-H2 difference, but there was considerable subject-to-subject variability within each group, with average F0-H2 across sentences ranging from 8.4 to 17.1dB in females, and from 4.6 to 9.7 in males. Periodicity versus noise excitation of F3 was measured for the reiterant sentences with [hɑ], on a subjective five-point scale and noise was found to be commonly present for both sexes with on average more noise in female than male subjects, but again considerable within-group variation. Both reiterant imitations of one of the original sentences pronounced by all subjects were then played to a panel of eight listeners, who were asked to judge the vowels on a seven-point scale from 'not breathy' to 'strongly breathy'. On average, females were perceived to be slightly more breathy than males, and sentences consisting of [hɑ] syllables were generally perceived as considerably more breathy than those with [ʔɑ]. Correlations of breathiness ratings with acoustic measures suggested that both the F0-H2 measure and the presence of noise were important. Finally, pairs of synthetic 'female' vowels (the first of each pair being a constant reference vowel) were played to a panel of five listeners who were asked to judge the relative breathiness of the second, its naturalness and its nasality. The results suggested that noise amplitude was more important than F0-H2 difference in giving a breathy percept; the latter cue was insufficient on its own to induce a breathy percept and often contributed to a perceived increase in nasality. The tentative conclusion of the authors is that,

> '... either breathiness is signalled differently for men and women, or that the increases in the first harmonic observed in production data from women must be accompanied by other cues to be interpreted by the listener as cues to breathiness.' (851)

Ní Chasaide and Gobl have published several papers developing the theme of the 1988 presentation mentioned above, among them one in *Speech Communication* (Gobl and Ní Chasaide 1992) where they analysed repetitions of a prose passage read with a range of voice qualities by a male phonetician who is a native speaker of British English. The data were subjected to manual interactive inverse filtering and analysed using the four-parameter LF-model of differentiated glottal flow developed by Gunnar Fant. Correlates of breathy voice were found to be high values for the parameters RA (corresponding to attenuation of higher frequencies), RK (corresponding to a more symmetrical pulse shape) and OQ (Open Quotient, thus also suggesting a more symmetrical pulse). Gobl and Ní Chasaide also used data from frequency domain analysis of the speech waveform to measure the levels of F1 and F2 relative to the first harmonic (our F0) and their Figure 5 (487) shows marked attenuation of both in the breathy data. An important feature to note about both sets of measurements is that they vary over time, and in their conclusion the authors emphasise the point that, 'a switch between voice qualities may not necessarily involve a single transformation which remains uniform throughout an utterance.'

Ní Chasaide and Gobl (1993) investigated voice quality in the vicinity of *Voiced* and *Voiceless* stop consonants spoken by male and female speakers in different languages. They found considerable cross-linguistic differences, but the effects were not grouped according to language-family as they had expected. Thus Swedish and, to a somewhat lesser extent, Italian /pː/ was preceded by a markedly higher RA than /b/, whereas, although the values were occasionally slightly higher in French and German (suggesting a slight tendency to relax the vocal folds in anticipation of the following *Voiceless* stop), the effect was not found to be consistent. The English speakers produced both patterns, but information is not given as to whether the division corresponds to the speaker's sex. RK values also rose in Swedish in anticipation of /pː/. Spectral measurements on the whole confirmed these findings, with the voicing category of the following consonant having little differential effect on F1 (their L1) relative to F0 in French and German, but showing a marked relative decline in F1 before the Swedish /pː/ with a rather lesser effect in the same direction in Italian. The English subjects fell into two groups, as for the source parameter measures. It is noticeable that for both sets of measures, the Figures show some marked differences between the languages, even within one of the two groupings (i.e. those with a /_p/ - /_b/ difference and those without).

In postconsonantal vowels, little categorial effect was found in the source parameters in French and Italian, but German RA was much higher at vowel onset following /p/ than /b/, and declined less rapidly. The authors infer that this is the result of incomplete glottal closure with the vocal folds vibrating in breathy mode following the aspirated stop. However, the difference between voicing categories is less marked in Swedish and English, despite the fact that these languages also have a voiceless unaspirated vs. voiceless aspirated phonetic contrast. The spectral data show less similarity between Swedish and the two Romance languages, with a lower F1 in Swedish post /p/ onset than following /b/, but no consistent effect in French or Italian. German follows a similar pattern to Swedish, but with an even greater relative lowering of F1. Data for English are not given. In the light of these findings, it is perhaps not surprising that no difference was found in the study reported above for vowels following voice versus voiceless stops in French.

A smaller-scale study is currently being carried out by Scobbie (1995 and personal communication), in which he found a marked difference between F0-H2 measures in vowel onset following /t/ vs. /d/, and to a lesser extent /p/ vs. /b/ in four-year-old speech-disordered child speakers of Edinburgh English.

5. Discussion

The 1988 study reported above raised several issues, to which we shall now return in the light of the subsequent work reported above.

5.1 Methodology

There are various methodological questions raised by a comparison of the studies mentioned, principal among which are how the oft-referred-to, but ill-defined feature 'spectral tilt' or 'spectral slope' is measured, and how measurements are analysed.

5.1.1 The measurement of spectral tilt.

The studies take one of two approaches to gaining access to an accurate measure of the voice source. Some invoke some procedure for negating the effects of the supra-glottal filter. Thus, Fant and Ní Chasaide and

Gobl used inverse filtering techniques, whereas Monsen and Engebretson had their subjects phonate down a reflectionless tube to reduce the resonances of the vocal tract. Bickley also used inverse filtering when she was looking at waveforms. The rest rely for the most part on analysing vowels with a relatively high F1 to minimise its effect on the lower harmonics, and/or on averaging large amounts of data to derive an accurate picture of the shape of the source spectrum. Henton and Bladon and Temple use statistical tests, while Hammarberg uses Long-Term Average Spectra (LTAS). Of course, with either approach it is impossible to be absolutely sure that a true picture of the glottal wave has been revealed, although inverse filtering techniques have improved greatly over recent years. The second type of approach seems the less satisfactory one, particularly for the purposes of comparing across studies, or even comparing different groups of speakers within studies: it is well-known that vowel qualities differ somewhat across languages (thus /a/ could represent something different in Gujarati from French), and across sex groups (and that the degree of sex-specific variation varies from language to language - see Bladon *et al* 1984)[13]. The fact that the trajectory for /a/ from position D to V in Figure 5 (above) is different from those of the other three vowels does suggest that we might be able to claim that the F1 transition is not affecting H2 in this case, but the uncomfortable fact remains that it is only this vowel which shows the unexpectedly steeper male slope in two positions. Moreover, Table 7 shows that only in a few measurements were the slope measurements for /a/ seen to be significantly different from those for the other vowels, where F1 is likely to have had an effect.

The actual measure of spectral tilt also differed from study to study. Fant and Ní Chasaide and Gobl used the LF model of glottal flow developed by the former, and measured parameters assumed to correspond to characteristics of the glottal wave. Because Hammarberg used LTAS, she was unable to make detailed measurements of spectral features, and instead identified breathy voice quality with relatively low energy in the F1 region (400-600Hz) and high levels in the highest frequency band (5-10kHz). Monsen and Engebretson measured slope in the first two octaves of their spectra in terms of dB fall-off per octave. Others measure formants, but in different ways: Barry compared amplitude levels for the same formant in his female and male subjects, while Gobl and Ní Chasaide measured F1 and F2 relative to F0. The rest of the studies measured harmonics, and I shall return to them in the next paragraph. The point needs to be made, however, that while these different measures allow generalised comparisons to be made of greater or less spectral tilt, the kind of detailed comparisons made, for example, between Henton and Bladon's data and that of Bickley is not possible.

---

[13] It could also be the case that /a̤/ and /a/ in Gujarati do not have the same formant values.

The studies using F0-H2 all measured the difference in amplitude between the two harmonics in dB. As we have seen, comparison using this measure between speakers with the same F0 is unproblematic (which is not to say that the interpretation of comparisons is without problems), but as soon as speakers with different F0 are compared, the analyst is faced with a choice which has implications for the results and can affect their statistical significance. Tables 10 and 11 present recalculations of Bickley's and Henton and Bladon's figures to see how this might affect the comparison between their sets of data.

	Difference (in dB/Hz)	
	Breathy	Clear
Speaker 1	0.1182	0
Speaker 2	-0.0364	-0.0273
Speaker 3	0.0182	-0.0273
Speaker 4	0.0455	-0.0364
Speaker 5	0.0455	-0.0818
Speaker 6	0.0364	-0.0727
Speaker 7	0.1	0
Speaker 8	0.0818	-0.0182
Speaker 9	0.1364	-0.0182
Speaker 10	0.0909	0.0182

Table 10. Slope between first and second harmonics for breathy and clear vowels (in dB/Hz) in !Xhóõ. Calculated from figures given in Table 1 above, assuming F0 to be 110 Hz.

Since the frequency data were not available, hypothetical values of 110 Hz for male speakers and 220 Hz for females were assumed. Moreover, only mean amplitude differences are available for Henton and Bladon's data.The Tables are intended to give an idea of how a different method of calculation might affect the comparison between them, rather than a mathematically precise reformulation of the data.

Vowel	/a/	/ɑ/	/ʌ/	//
Females	0.0382	0.0291	0.0282	0.0150
Males	0.0089	0.0070	0.0015	0.0036

Table 11. Average slope (in dB/Hz) between the first and second harmonics in male and female speakers of Received Pronunciation. Calculated from figures given in Table 3 above, assuming F0 to be 220 Hz for female speakers and 110 Hz for male

Table 11 shows a clear difference still between the male and female RP speakers and the female slopes are still steeper than the !Xhóõ clear vowels. However, whereas the F0-H2 amplitude difference for the RP females' /a/, /ɑ/ and /ʌ/ was greater than for six of the !Xhóõ breathy vowels, it is only greater than two in the dB/Hz measure (with /a/ alone being greater than one other in addition). Moreover, if the RP female /a/ measurement is compared with, for example, !Xhóõ speaker 10, the ratio is 0.84 on the dB measure, but only 0.42 on the slope measure. More significantly, the recalculation changes the relationship of the measurements of the RP speakers with the evaluations of Bickley's phoneticians. The recalculated average amplitude differences for vowels judged to be in the four categories of breathiness (see p.4 above for dB figures) are as follows: 'Very breathy' - 0.1136 dB/Hz, 0.0909 dB/Hz; 'Breathy' - 0.0755 dB/Hz, 0.1 dB/Hz; 'Slightly breathy' - 0.0609 dB/Hz, 0.0482; 'Not breathy' 0dB/Hz, 0 dB/hZ. When these values are compared with the RP females, the latter are seen not even to reach the 'Slightly breathy' level. It is the case that many of the Gujarati and !Xhóõ vowels also do not reach that level in either measure, and it must be remembered that the phoneticians were asked to judge degree of breathiness rather than whether the vowels were breathy or not, and that these are average values. Nevertheless, these calculations show that there are potential problems for comparative statements which remain to be resolved.

It is evident that further experiments are needed to test whether the straightforward amplitude difference between successive harmonics, or the 'slope' between them is perceptually salient. The evidence reviewed in the present article provides little basis for deciding between the measures, but Monsen and Engebretson's suggestion that there is some sort of built-in normalisation factor in the differing slopes (see Fig. 1 and comments in section 2.2 above) would imply that maybe it is the slope which is important. Figure 1(b) shows the near-identity of the spectral envelopes in un-normalised spectra: it is not the amplitude difference alone between each pair of harmonics which allows this to happen, but the combined effect of that and the distance between them in frequency.

5.1.2 The use of statistics

Many of the studies discussed, use statistical analyses of the data. This not only poses problems of comparability between studies because of the different numbers of subjects studied, but also those studies which present only statistical comparisons of groups of speakers risk masking variability within each group. Dempster (1992) illustrates this dramatically with an analysis of F0-H2 differences in two contexts in the large DARPA TIMIT Acoustic-Phonetic Speech Database Training Set, a database containing material from 420 speakers of U.S. English. Whilst one might want to take issue with aspects of Dempster's study, his evidence for the dangers of relying on statistics for drawing conclusions is salutary: he found a statistically significant difference ($p<0.1$) between male and female F0-H2 differences for the vowel /aa/[14] (measured in dB), but when the data are presented in histogram form, a very large degree of overlap is apparent.

While it is right, as Dempster says, that we should heed Klatt and Klatt's warning that, 'it is unwise to make sweeping generalisations with regard to sex typing' (*op. cit* 852), this does not invalidate or preclude further exploration of some of the questions raised in the present paper concerning the undoubtedly strong sex-specific tendency found in the work reviewed.

---

[14] TIMIT phonetic label representing the vowel in *heart* etc.

5.1.3 Perceptual experiments

All the perceptual experiments reported involve trained phoneticians. The answers thus tell us whether phoneticians judge the voice qualities according to a linear scale of 'breathiness' which they have learned. This does not really tease out the different contributing factors or enable us to make much progress with one of the central questions, that is whether the findings discussed above are addressing something which can really be construed as the same phenomenon in the real world. For example, does F0-H2 difference contribute to the perception of [a] versus [a̤] for the ordinary, untrained speaker of Gujarati?

That the judgements elicited tend to be on a scale of breathiness is also worthy of comment. When breathiness is being examined as a possible correlate of maleness or femaleness, or of degree of severity of a pathological condition, the justification for the approach is evident, but in an investigation of the acoustic correlates of phonological categories its relevance is less clear (compare, for example, the fact that English native speakers do not tend to hear absolute initial prevoiced French stops as 'very voiced'; when students of French are asked to attend to prevoicing, they often perceive a preconsonantal nasal element.)

5.2 Are we all talking about the same thing?

Perhaps the most important question, and one which needs to be considered before further detailed investigations of some of the problems highlighted in this paper are carried out,, is whether we are not being mislead by applying a single label to a variety of phenomena which are different in some respects. There is common ground between all the studies discussed, but they are looking at spectral tilt as a marker of breathiness in four different contexts:

1. as indicative of male-female physiological differences (e.g. Monsen and Engebretson);

2. as indicative of breathy voice quality for sociolinguistic or paralinguistic effect (e.g. Henton and Bladon);
3. as a characteristic of phonological categories (e.g. Bickley);
4. as indicative of a pathological problem (e.g. Hammarberg).

Is it justifiable to extend Ladefoged's 1983 statement quoted earlier to apply to the studies reviewed here? That is, is it really reasonable to claim that the 'breathiness' of pathological subjects or Gujarati speakers' [a̤] vowels, rather than a tendency for the difference between F0 and H2 to be greater, is characteristic of female speech? Barry's finding that noise in the high-frequency regions of the spectrum was as important for generating a 'good match' female voice suggests that it may be, and indeed the vibratory pattern suggested by Monsen and Engebretson for female vocal folds would predict that more noise would be generated than by males, as well as females having an enhanced fundamental. But this does not guarantee that the relative 'amounts' of noise and tilt are the same in all the cases. If, as Klatt and Klatt claim, noise is more important than tilt for giving a breathy percept, then maybe the F0-H2 differences found by Henton and Bladon are not indicative of breathiness at all.

In addition, the physiological correlates of the acoustic phenomena are reported or hypothesised to be different in the different cases: Ladefoged (see page 2 above) describes different correlates for breathiness in Gujarati vowels and English voiced /h/, the former a deliberate configuration of the vocal folds, and the latter a passive effect; Hammarberg posits incomplete abduction of the vocal folds as a result of unilateral paralysis or nodules on the folds; and Monsen and Engebretson ascribe the greater spectral tilt and noise to the different vibratory patterns of the vocal folds in males and females, which are in turn caused by differences in mass and structure. There is no reason why the relationship between production settings and acoustic structure has to be one-to-one, but it cannot be taken for granted that the different settings will necessarily produce something which can be called the same.

REFERENCES

Barry, M. 1986. Synthesising female voice quality: parameters and test methods. *Cambridge Papers in Phonetics and Experimental Linguistics* 5.

Bladon, R. A. W., Henton, C. G. and Pickering, J. B. 1984. Outline of an auditory theory of speaker normalization. In van den Broecke, M. P. R. and Cohen, M. (eds) *Proceedings of the Xth International Congress of Phonetic Sciences* 313-317.

Bickley, C. 1982. Acoustic analysis and perception of breathy vowels. *MIT Working Papers in Speech Communication* 1. 71-81.

Ní Chasaide, A. and Gobl, C. 1988. Voicing contrasts and the voice source. Paper delivered at the British Association of Academic Phoneticians Colloquium, Dublin.

Ní Chasaide, A. and Gobl, C. 1993. Contextual variation of the vowel voice source as a function of adjacent consonants. *Language and Speech* 36. 303-330.

Clark, C. J. 1986. Description of a speech analysis system. *Progress Reports from Oxford PHonetics* 1. 13-25.

Dempster, G. J. 1992. Acoustic cues to breathiness: a true marker of gender? *Proceedings of the Institute of Acoustics*. Vol. 14, Part 6. 249-256.

Gobl, C. and Ní Chasaide, A. 1992. Acoustic characteristics of voice quality. *Speech Communication* 11. 481-490.

Hammarberg, B. 1986. *Perceptual and Acoustic Analysis of Dysphonia*. Dissertation from Dept of Logopedics and Phoniatrics, Huddinge University Hospital, Sweden.

Henton, C. G. and Bladon, R. A. W. 1985. Breathiness in normal female speech: inefficiency versus desirability. *Language and Communication* 5. 221-227.

Klatt, D. H. 1986. Detailed spectral analysis of a female voice. *Journal of the Acoustical Society of America* 80, Supplement 1. S69.

Klatt, D. H. and Klatt, L. C. 1990. Analysis, synthesis and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America* 87. 820-857.

Ladefoged, P. 1981. The relative nature of voice quality. *Journal of the Acoustical Society of America* 69, Supplement 1. DD3.

Ladefoged, P. 1982. *A Course in Phonetics*. New York: Harcourt, Brace Jovanovich.

Ladefoged, P. 1983. Linguistic uses of different phonation types. In Bless, D. M. and Abbs, J. H. (eds). *Vocal Fold Physiology*. San Diego: College Hill Press. 351-360.

Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge: C.U.P.

Mattingly, I. G. 1966. Speaker variation and vocal tract size. *Journal of the Acoustical Society of America* 39. 1219.

Monsen, R. and Engebretson, A. M. 1977. Study of variations in the male and female glottal wave. *Journal of the Acoustical Society of America*. 62. 981-993.

Ohala, J. J. 1983. Cross-language use of pitch: an ethological view. *Phonetica*. 40. 1-18.

Sachs, J. P. 1975. Cues to the identification of sex in children's speech. In Thorne, B. and Henley, N. (eds) *Language and Sex: difference and dominance*. 152-171.

Scobbie, J. M. 1995. Phonological and phonetic perspectives on the delayed and disordered acquisition of English initial consonant clusters. Paper presented at the Third Phonology Workshop, North-West Centre for Romance Linguistics, Manchester, May, 1995.

Temple, R. A. M. 1988a. *Sex-specific Aspects of the Voicing Contrast in French Stop Consonants*. Oxford M.Phil. thesis.

Temple, R. A. M. 1988b. In search of sex-specific differences in the voicing of French stop consonants. *Progress Reports from Oxford Phonetics* 3. 74-99.

Watson, I. M. C. 1987. Problems in quantifying the child voicing contrast. Ms, University of Oxford.

NOTES ON TEMPORAL INTERPRETATION AND CONTROL IN MODERN GREEK GERUNDS[*]

George Tsoulas

Department of Language and Linguistic Science
University of York

1. Introduction

In this paper I would like to examine some aspects of the syntax of the Modern Greek gerund clauses. This study will mainly focus on the following aspects of the syntax of these clausal constituents:

(i) Their External and Internal Syntax
(ii) Temporal Interpretation of Gerund clauses
(iii) Their Argument status
(iv) Control in Gerunds

As a starting point in this paper we adopt the commonly held view that gerund clauses are never arguments but only adjunct modifiers. Our account of their temporal interpretation relies on recent theories of adjunction under which the configurational difference between adjuncts

---

[*] Earlier versions of this paper have been presented to the first Workshop on Modern Greek Syntax in Berlin on December 1994 and at the CNRS in Paris (URA 1720) on February 1995. I want to thank these audiences for their comments and discussion. Particularly I would like to thank Artemis Alexiadou, Sabine Iatridou, Lea Nash, Alain Rouveret, Anne Zribi-Hertz. Thanks also to David Adger for very useful comments and discussion on a preliminary version of this work. Needless to say I am alone responsible for the views defended here as well as for all remaining errors of fact and interpretation.

York Papers in Linguistics 17 (1996) 441-470
© George Tsoulas

and specifiers vanishes. Furthermore, we provide arguments from ECM constructions, imperatives and topicalisation in favour of the claim that gerund clauses can also be arguments. This in turn leads us to a principled account of the puzzling control patterns found in gerund clauses.

2. An Overview of the Issues

Consider the following Modern Greek sentences:

(1) I Maria$_i$ ide to Gianni$_j$ [$_{CP}$ PRO$_{*i/j}$ zografizondas ena dendro].
The Maria saw the Gianni painting a tree
Maria saw Gianni while he was painting a tree.

(2) I Maria$_i$ ide to Gianni$_j$ [$_{CP}$ PRO$_{i/*j}$ zografizondas to dendro].
The Maria saw the Gianni painting the tree
Maria saw Gianni while she was painting the tree.

Under currently quite standard assumptions concerning the nature and the sites of adjunction (Chomsky 1989, 1992, 1993; Kayne 1994) one may suppose that there is no significant structural difference in the syntax of sentences (1) and (2). As the indexing indicates however there is a difference in so far as the controller of the PRO is concerned. The only observable difference in the two sentences is the nature of the object of the verbal form *zografizondas*: in (1) the object of this verb[1] is an indefinite DP, and in (2) it is a definite DP.

Notice also that in a sentence like (3), in which (2) is embedded under the verb *Akousa* 'I heard', the controller cannot be the subject of the main clause (*pro* with first person features).

(3) Akousa oti i Maria ide to Gianni zografizondas to dendro.
Heard/I that the M saw the G painting the tree

---

[1] Although the precise nature of this form remains to be determined we will use *verb* for the moment for convenience.

Bearing in mind that the gerund, as the glosses indicate, has a specific temporal interpretation, one question that we have to address is why in (3) the gerund clause cannot be associated with the matrix.

A further issue arising is whether the object *To Gianni*, which displays accusative Case, genuinely belongs to the matrix sentence or whether it is in fact the subject of the gerund clause which is Exceptionally Case Marked by the higher verb. In order to provide a satisfactory answer to this question one has to settle the issue of the argument status of the gerund clause.

As will become clear in the remainder of the paper the differences seen above in syntax and interpretation are due to the ambiguity of these forms, which can be either participles or gerunds. The paper is organised as follows. In the following section I present the distribution of gerund clauses. Then I examine their categorial status and their internal syntax, focusing principally on their temporal interpretation and several temporal scope ambiguities. In the last part I examine their argument status and modify the initial assumption that gerunds in Modern Greek are only adjunct modifiers. I conclude with a discussion of the control properties of gerunds.

3. The Modern Greek Gerund

In this section I want to investigate the properties of what has been frequently called a gerund in Modern Greek. This form is exemplified in (4).

(4) Pinondas to krasi
drinking the wine

This verbal form has not received much attention in the recent literature.[2] The question of what its precise nature is and its place within the Modern Greek verbal paradigm has not yet been clearly addressed. In fact whenever, in the literature, (4) is put under the heading *gerund*, it is only because of its apparent lack of agreement and tense features.[3] On the other hand, the fact that this form, historically, clearly derives from the active participle has led some researchers to classify it with participles. In this paper I will argue that this form is ambiguous in that in some cases it behaves as a participle, and in others more as a gerund. Two caveats are in order here. First, as will become apparent in the remainder of this paper, it would be misleading to understand by the term *gerund* the notoriously syntactically and semantically ambiguous English counterpart. Only one aspect of the function and distribution of the English gerund is displayed by the Modern Greek (4). Examples (8)-(11) are intended to show this.

---

[2] Not only in recent years but also in the literature since the 1930s, to the best of my knowledge, this form received only a passing mention in the morphology section of reference grammars and other works. Its syntax has never really been seriously investigated, see for example Joseph and Philippaki-Warburton 1986, Householder, Kazazis and Koutsoudas 1964, Tzartzanos 1949, Seiler 1952, Mirambel 1939 among others.

Second, the participial uses of (4) are not on a par with the uses of clearly participial forms in Modern Greek: although the gerund can be considered a participle in so far as it restricts the possibilities of control, it still preserves other verbal properties whereas real participles do not.

Examples (5)-(11) cover essentially the distribution of the Modern Greek gerund.

(5) **Pinondas** to krasi o Giannis kapnize.
drinking the wine the Giannis was smoking
Giannis was smoking while he was drinking the wine.

(6) O Kostas kimotan **kratondas** to molyvi tou.
The Kostas was sleeping holding the pen his
Kostas was sleeping holding his pen (with his pen in his hand).

(7) **Rixnondas** to potiri to espase.
dropping the glass it (S)he broke
She broke the glass by dropping it.

---

[3] With the notable exception of Householder, Kazazis and Koutsoudas 1964 who provide more evidence for such a claim (see below).

(8) * O Giannis ekseplagi apo to **telionondas** tou arthrou.
The G. was surprised by the finishing of the paper
Finishing the paper was a fact that surprised Giannis.

(9) * O Kostas pige **psarevondas.**[4]
The Kostas went fishing
Kostas went fishing.

(10) * (To) **telionondas** to arthro toso grigora mas ekseplikse.
(The) finishing the paper so quickly us surprised
Finishing the paper so quickly was a fact that surprised us.

(11) * O kostas zitise **arcizondas** mathimata pianou.
The Kostas asked starting lessons piano
Kostas asked to start taking up piano lessons.

It is clear from the above examples that gerundival clauses only appear as adjunct modifiers (5, 6, 7), they can never be subjects or objects of verbs or prepositions (8, 9, 10, 11); they can never occupy an A-position. They can however be adjoined to various sites depending on their meaning and in that respect they are parallel to adverbial modifiers. Thus, a manner gerund will be adjoined to VP, a temporal gerund is adjoined to IP and a modal even higher, as in (14).

(12) I Anna anisixise to Niko **fonazondas** voithia.
The Anna worried the Niko crying out help
Anna caused worry to Noko when (because) she cried out for help.

(13) I Anna **ftiaxnondas** kafe milai sto tilefono.
The Anna fixing coffee (she)speaks on the phone
Anna talks on the phone while she is making coffee.

---

[4] I leave aside here the idiomatic *pigeno girevondas* 'I am looking for trouble'.

(14) **Echondas** makria malia i Anna prepi na ta xtenizi sinechia.
Having long hair the A. must C them comb always
Having long hair Anna must comb it all the time.

This difference in the semantic interpretation as reflected by the syntax can be explained by a difference in intensionality. In (12) one may suppose that given that the contents of the VP have all moved higher to functional projections the gerund remains adjoined to the VP. In (13) the subject is outside the scope of the adjunct but the remainder of the VP is not. In (14) the gerund has in its scope something akin to the Σ Phrase of Laka (1990) which explains its modal interpretation.

2.1 External Distribution[5]

What I call here gerund has frequently been confused with participles and, consequently, it has been considered a 'nominal' form of the verb. However there is clear evidence that the gerund shares distribution with verbs. Gerunds are opposed to participles in that they can never be nominalised (see (15)), i.e. they can never be preceded by a determiner; they can only be modified by adverbs (see (16) and (17)); they do not compose with auxiliaries to form complex tenses (see (18)); and, in general, they only function as verbs. Participles, on the other hand have all the opposite properties, (except for the complex tenses[6]) as the following examples show.

---

[5] I am interested here in the overall behaviour of the gerund and not in its precise morphological constitution. Due to space limitations I will not attempt here to analyse the function of the morpheme *-ondas* that forms the gerund. Historically, this morpheme comes from the accusative of the active participle of Ancient Greek (with the rather mysterious addition of the -s ending). I believe that this resemblance and historical affiliation is responsible for much of the confusion created among scholars as to the nature of the gerund. I leave a more detailed analysis of its morphological peculiarities for further research.

[6] Strictly speaking participles do not either compose with auxiliaries to form complex tenses. Complex Tenses in Modern Greek are formed by means of a different form, derived from the past tense's root together with a third person singular ending (with some exceptions), this form is not homophonous to the third person singular of the past tense because it lacks the temporal prefix (augment) /e/. However, the investigation of the

GERUNDS

(15) * To kskinondas ine diskolo.
The starting is difficult

(16) * To ksekinondas, to opio theloume[7] ine diskolo.
The starting the which (we) want is difficult

(17) Milouse kitondas me astamatita.
he/she was talking looking at me all the time

(18) * echo/ime kitondas.
I have/be looking

PARTICIPLES

(19) O Xaroumenos ine efxaristos.
The happy/MASC is pleasant

(20) O Xaroumenos anthropos ine efxaristos.
The happy/MASC man is pleasant

(21) O Xamenos, o opios bori na ine opiosdipote, den xerete.
the looser/M the which can C be anyone neg rejoice

(22) Milouse arnoumeni na me kitaksi.
she was talking refusing/F C at me look
She was talking refusing to look at me.

---

morphological properties of this form would take us too far astray from our initial purposes. I will thus leave it aside for the present paper.

[7] Here the modifier is a relative clause. Examples showing the gerund being modified by an adjective are not particularly illuminating since the gerund, uninflected for gender, would have to be modified by a third person neuter adjective, a form which, in Modern Greek, coincides with the adverb. Notice also that in (16) the presence (or absence) of the determiner *To* is irrelevant to the grammaticality of the sentence.

These examples show that the distribution of the gerund can be considered as a subset of the distribution of the participle. Participles are in principle categorially ambiguous in the sense that they can function either as verbs or as nouns or adjectives. The distribution of the gerund covers only one part, the verbal part, of the participle's distribution. Differently put, only example (22) is comparable to the examples (12)-(14) which show the distribution of gerunds.

3.2 The Structure of Gerund Clauses

The main question arising in connection with the internal structure of gerund clauses is their categorial status, this question will be shown to be of a major importance because it bears directly on the status of their subject. Gerund clauses seem to be CPs. In the following examples, cases of wh-extraction from within the gerund clause are shown.[8]

(23) Ti$_i$ pinondas akouge mousiki?
what drinking (s)he listening music
What was she drinking while she listened to the music?

(24) Se pion milondas magireve?
To whom talking he/she was cooking
Who was she talking to while she was cooking?

(25) Pou kitondas sou milouse?
where looking to you was talkng
Where was she looking while she was talking to you?

In (23) and (24) argument extraction is displayed (direct and indirect object respectively) and (24) shows adjunct extraction.[9] These examples show that a Spec, CP position is available and can be targeted by wh-movement. On the other hand, similar examples involving clearly participial forms (i.e. inflected for number, gender, person, and Case) are sharply ungrammatical:

---

[8] All the sentences involving extraction are somehow marginal in acceptability. Their marginal status is to be imputed to the well known fact that extraction out of an adjunct is generally marginal. The relevance of these examples will become more evident when they are compared with extraction out of participles, which is impossible.

[9] There is of course the possibility of leaving the wh in situ, which is also more natural (but see note 8):

(i) Pinondas ti akouge mousiki

(26) * Ti ton thimasai arnoumeno.
what him remember/you refusing/3/S/M/ACC

(27) * Pou ton ides vriskomeno.
where him saw/you being/M/S/3/Acc

(28) ??Pou ton ides eksaskoumeno?
where him you saw exercising
Where did you see him exercising?

There is a difference in acceptability between (26)-(27) and (28) which is much better. The reason for this asymmetry between argument/adjunct extraction is obscure. Notice that the locative in (27) behaves more like an argument of the verb *vriskomai* 'being in a location'.[10]

These examples suggest that, contrary to gerunds, participial clauses are bare IPs (or even VPs). This observation is particularly significant for the subpart of the distribution of participles that coincides with the distribution of gerunds, i.e. when participles function as verbs.[11]

---

drinking what was/(s)he listening to the music
(ii) Milondas se pion magireve
talking to whom was/(s)he cooking
(iii) kitondas pou sou milouse
looking where to you was (s)he talking
(S)he was talking to you looking where?

[10] This type of asymmetries suggests that the lexical semantics of each item have some influence, but I will not pursue this path further.

[11] It is rather interesting to note that for some obscure reason the option of long wh-movement, widely attested in pro-drop languages such as Modern Greek, is not available here.

3.2.1 Temporal Interpretation of Gerunds

Gerunds are further opposed to participles in that, aspectually, they are uniformly imperfectives whereas participles are perfectives.

(29) Pinondas arga to krasi milouse gia glossologia.
drinking slowly the wine he was talking about linguistics

(30) diavaze kapnizondas astamatita.
he was reading smoking without stopping

(31) * Arnoumenos arga tin prosfora efige.
Refusing/3/S/M/Nom slowly the offer left/he

(32) Eksaskoumenos astamatita katafere to skopo tou.
exercising/MASC all the time he reached the aim his

The perfective/imperfective difference can also be cast in terms of definiteness/indefiniteness. I have proposed in Tsoulas (1994a, 1994b, 1995) that tense is also subject to the definiteness/indefiniteness distinction. Furthermore, I have proposed that this distinction should replace the classical finite/non-finite distinction, since it is now widely accepted that non-finite verbal forms only lack morphological temporal specifications, while semantically still they contain information pertaining to temporal interpretation. This theory has interesting predictions in that it parallels clausal and nominal (DP) constituents in yet one more respect. Informally in the case under examination, the gerund is indefinite in that it does not refer to a precise point or interval in time whereas participles do In the grammatical example (32) the temporal reference of the participle can be characterised as a closed temporal interval located at some time before the occurrence of the event denoted by the main verb. By contrast, gerunds denote open intervals with respect to the main verb. If we consider gerunds as indefinites, this constitutes an additional explanation for the extraction data in the preceding paragraph, namely, indefinites permit extraction while definites disallow it (see Ross 1968, Manzini 1993 among others).

3.2.2 Temporal Scope Ambiguities with Gerund Clauses

In this subsection I will present some more evidence for the CP status of gerund clauses. This evidence also bears on the issues of control mentioned in the introduction. This evidence involves temporal scope ambiguities and binding with gerunds. Consider the following sentences:

(33) Tremondas apo to fovo tou o Giannis lei oti o Kostas efige.
trembling by the fear his the G. says that the K. left
Giannis says that Kostas left trembling from fear.

(34) Vlepondas ta ligosta malia tou o Giannis ipe
Seeing the few hair his the G said
oti o Kostas epathe egefaliko.
that the K. had a stroke
Giannis said that Kostas had a stroke seeing his thining hair.

(35) Trogondas ti soupa tou o Giannis ipe oti o Kostas kaike.
Eating the soup his the G said that the K. was burned
Giannis said that Kostas burned himself while eating his soup.

(36) Ida to Gianni vgainondas apo to spiti (tou)
Saw/I the G. coming out of the house (his)
prin na ton skotosi o Kostas.
before C him killed the K
I saw G. getting out of his/the house before K. killed him.

(37) Ida ton Kosta na skotoni to Gianni vgainondas apo to spiti (tou).
Saw/I the K C kill the G coming out of the house (his)
I saw Kostas killing Giannis while getting out of the/his house.

(38) Ematha oti o Kostas skotose to Gianni vgainondas apo
Learned/I that the K. killed the G coming out of
to spiti tou prin mathefti o tsakomos tous.
the house his before becomes-known the fight their
I learned that K killed G getting out of the/his house before
their fight becomes known.

Examples (33)-(38) show that the gerund can be construed with each of the clauses in the complex structure. For example, (38) can have the following interpretations:

(i) I heard, when I was getting out of the/his house that Kostas killed Gianni, before their fight becomes known.

(ii) I heard that Kostas, as he (Kostas) was getting out of the/his house he (Kostas) killed Gianni, before their fight becomes known.

(iii) I heard that Kostas killed Gianni when he (Gianni) was getting out of the/his house before their fight becomes known.

Interestingly enough the gerund clause cannot be associated with the *before*-clause in this structure. We will be merely noting this fact for the moment, we shall return to it shortly.

In general, it is natural to suppose that the adjunction site is what determines the interpretation. In other words, the gerund clause must be adjoined to a given T (or I) node in order to be able to modify that node. However, we see that the same surface string can yield several interpretations. The question is how these interpretations are to be derived in a framework like the minimalist program (Chomsky 1993, 1994, 1995), where one of the major predictions of the theory is that optionality should be banned. One way to deal with this problem is to suppose that the entire adjunct is covertly moved and readjoined to some other position. One may, however, legitimately ask what motivates such a movement, since all movement operations must be driven by the need to check some morphological feature. It is difficult to imagine what that feature could be. Another way around this problem that comes to mind derives from Geis' (Geis 1970) and Larson's treatment of temporal prepositions as involving silent temporal operators that need to be moved to the COMP position of the clausal complement of the preposition.[12] Consider for example a sentence containing a *before*-clause:

---

[12] Cited by Johnson 1988, who applies this analysis to clausal gerunds in English.

(39) Valerie arrived before you said she had.[13]

This sentence is ambiguous. It has one meaning corresponding to (i) and one meaning corresponding to (ii).

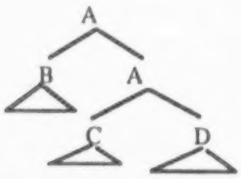
(i) Valerie left before the time of your saying that she had.
(ii) Valerie left before the time you said she had left at.

According to Larson, as cited by Johnson (1988), the ambiguity arises because in these clauses there are empty temporal operators. These operators, once moved to the appropriate position, bind a variable located either in the matrix (i) or in the embedded clause (ii). This analysis, since it is based on movement, has the major prediction, as noted by Larson and Johnson, that the interpretation of this type of sentences would be sensitive to island effects (see Johnson 1988 for the relevant examples and discussion). This prediction, which is indeed a true one, raises a major problem for the syntax of Modern Greek gerunds. If we assume that a similar analysis can be proposed for gerunds in Modern Greek then movement of the operator out of the adjunct would violate the adjunct condition and yield ungrammatical results. In the examples (33)- (38) the gerund always has scope over one of the clauses in the structure excluding all the others. This fact is an argument in favour of the analysis in terms of movement of a covert operator in the sense that it makes it necessary to understand scope in this particular context as the relation between an operator and the variable it is associated with (i.e. that it binds), rather than in terms of C-command or any other command-type relation. This fact is of a crucial importance given the theory of adjunction we are adopting in this work, to which I turn in a moment. Suppose that this analysis is correct and Modern Greek gerunds truly contain a phonologically null temporal operator (a silent *when or while*); how can we account for the improper movement of the operator out of the gerund? In order to answer this question let us turn first to the nature of structures formed by adjunction. Kayne (1994) proposes that there is no principled difference between a specifier and an adjoined element, under this assumption and given a phrase marker like (40) where B is adjoined to A, if B represents the gerund clause of our examples and A is, say, a VP or IP, then no locality problem arises if we move the operator to the first superordinate CP position.[14]

---

[13] Example adapted from Johnson 1988.

(40)



This type of movement requires that the B adjunct be a CP projection, for, otherwise the derivation would be ruled out as an ECP violation while here antecedent government is satisfied. It is also interesting to observe that even in (41) the gerund can still be associated with the matrix clause, in the interpretation that the *learning* event takes place when the *learner* steps out of her house.[15]

(41) Ematha oti o Kostas ipe oti o Nikos skotose to Gianni
Learned/I that the K. said that the N. killed the G.
vgainondas apo to spiti.
coming out of the house
I learned that Kostas said that Nikos killed Gianns while
getting out of the house.

If my analysis so far is correct we have to assume that only the operator itself can bind an event variable, and, crucially, not its trace

---

[14] Recall that we analyse gerunds as indefinites, thus allowing material from within the gerund clause to be extracted.

[15] Predictably, this reading is somewhat more difficult to obtain. It is noteworthy that, in general, speakers require a clear pause before the adjunct in that reading, this requirement is weakened though if the choice of lexical items is such that the association of the gerund with another clause is unlikely.

(t_{op}) since to satisfy the ECP the operator has to move stepwise through the specifiers of each of the embedded CPs. If t_{op} were to be a potential binder for the event variable of each verb, the whole structure would be uninterpretable and the derivation would crash as a violation of the bijection principle of Koopman and Sportiche(1984).[16]

Returning to our example (38), under this analysis this example should be problematic since under our assumption that there is no principled, configurational difference between adjuncts and specifiers, nothing would prevent the operator contained in the gerund clause from moving to the specifier of the clausal complement of the preposition *prin*. Recall however that the analysis proposed here crucially assumes that these temporal operators are also present in other temporal clauses, including *before*-clauses. Therefore it is impossible for the temporal operator of the gerund clause to move into the position that is already occupied by the operator originating in the *prin*-clause. Consequently in sentence (38) the only interpretation of the *prin*-clause with respect to the matrix is a narrow scope interpretation, which means that the time that *prin* 'before' compares can only be construed with one of the embedded clauses but crucially not with the gerund or the matrix clause.

3.2.3 Manner and Modal Gerunds

The analysis presented so far covers mainly temporal (and aspectual) gerunds. Manner gerunds behave in almost the same way. Consider (41) in a manner reading of the gerund. Suppose that (41) is uttered in order to describe a particular scene of a gang fight where Nikos killed Gianni as he (Nikos) was shooting his way out of the house. I propose that this interpretation will not be merely the result of the fact that the gerund is adjoined to the lowest VP but because the temporal operator will move to the Spec of the most deeply embedded CP and no further up. Strictly speaking, these should be considered as two relatively independent processes. For one thing, the gerund has a specific dependent temporal interpretation and this must somehow be accounted for.[17] Its adjunct status requires a different mechanism from those given in Tsoulas (1994a, 1994b) for the interpretation of indefinite clausal constituents. The data examined there involved, crucially, sentential complements. Thus, although the adjunction site is still crucial to the interpretation, it is the temporal operator that determines in a complex structure with respect to which such adjunction site the gerund clause will be interpreted.[18] Consider now (42) in which the gerund is clearly denoting manner:

---

[16] This is quite natural. The operator and its trace are non distinct under the copy theory of movement, since they share the same index.

(42) Ematha oti o Kostas ipe oti o Nikos skotose to Gianni
I learned that the K. said that the N. killed the/Acc G.
pirovolondas ton.
shooting him
I learned that Kostas said that Nikos killed Giannis shooting him.

---

[17] An Indefinite one as we said above. The morphological expression of the temporal indefiniteness in this case is quite a distinct matter. Along the lines of Tsoulas 1994a, if the generalisation concerning the morphological realisation of temporal indefiniteness, is correct, we infer from the existence of special bound morphology on the verb, that the [-DEFINITE] feature is realised under I (or T). This generalisation states that temporal (clausal) indefiniteness can either be realised in I or in C and either as bound morpholgy on the verb or as an independent word, moreover whenever temporal indefiniteness is realised as a bound morpheme it is necessarily realised under I. These facts, in conjunction with the ones about temporal indefiniteness in French presented in Tsoulas 1994a, b, 1995 raise a serious problem, namely, it shows quite clearly that the morphological realisation site, differing between C° and I (T) is not really subject to parametric variation since the two options exist within the same language, French as well as Modern Greek. The reasons for this optionally I don't really understand for the moment. They might have to do with the availability of control into the indefinite clausal constituent, but even this line of reasoning is compromised by the Modern Greek data, since in Modern Greek control is available both in subjunctives (Indefiniteness in C) and Gerunds (Indefiniteness in I). I will leave the matter here for this paper and postpone a more detailed examination for further research.

[18] Semantically this account is also supported because of its compositionality.

It could be objected that in this case the previous account somehow fails to capture the fact that the gerund can only be associated with the lowest VP. In a way, it is entailed by the lexical meaning of each item that the gerund says something about the manner in which the killing took place. This is not strictly true however, it is also conceivable that the clitic pronoun *ton* does not in fact refer to the DP *to Gianni* (the killed man) but rather it picks out some other antecedent from the preceding discourse. In this case, assuming for concreteness that the temporal operator has moved to the [Spec CP] of the matrix, the intended meaning is that the speaker learned about the facts reported *when* she was shooting someone. This becomes even clearer in (43).

(43) Akousa oti o Kostas ipe oti o Nikos skotose to Gianni
Heard/I that the K. said that the N. killed the G.
pirovolondas tin.
shooting her

The replacement of the masculine *ton* by a feminine form prevents its association with any of the DPs present in the sentence. (43) remains however grammatical, within, of course, the appropriate context.

The same considerations apply also to modal gerunds though the facts get somewhat more complicated in this case, for reasons I don't fully understand. Consider the following examples (partly adapted from Stump 1985). In this set of examples we show Modal gerundival clauses adjoined to various positions in the complex structures. Interestingly, the temporal patterns shown are not homogeneous. They differ in that the gerund clause in the examples (48)-(52) cannot be freely associated with any of the other clauses in the complex structure.

(44) forondas afta ta rouha trelene olo ton kosmo.
wearing these the clothes he/She was driving mad all the people
Wearing this outfit (s)he was driving everybody crazy.

(45) Akousa oti o Kostas ipe oti o Nikos itan sigouros oti
Heard/I that the K. said that the N. was sure that
forondas afta ta rouha tha trelenotan olos o kosmos.
wearing these the clothes would be driven mad all the people
I heard that Kostas said that that Nikos was sure that wearing
this outfit, he would drive everybody mad.

(46) Pernondas to farmako se kanoniki dosi,
Taking this drug in normal dose
vlepis grigora apotelesmata.
see/you quick results
You see prompt results if you take this drug in normal dose.

(47) Vlepis grigora apotelesmata,
See/you quick results
pernondas to farmako se kanoniki dosi.
taking this drug in normal dose
You see prompt results if you take this drug in normal dose.

(48) Akousa oti o Kostas ipe oti o Nikos itan sigouros oti
Heard/I that the K. said that the N. was sure that
Pernondas to farmako se kanoniki dosi,
taking the drug in normal dose
ta apotelesmata ine theamatika.
the results are spectacular
I heard that Kostas said that Nikos was sure that you see
prompt results if you take this drug in normal dose.

(49) Echondas makria heria o Nikos ftanei efkola to tavani.
Having long arms the N. reaches easily the ceiling
Having long arms Nikos reaches easily the ceiling

(50) *O Nikos ftanei efkola to tavani, echondas makria heria.
The N. reaches easily the ceiling, having long arms
Having long arms Nikos reaches easily the ceiling

(51) O Giannis kseri oti i Eleni ipe oti echondas makria heria
The G knows that the/fem E. said that having long
arms ftanei efkola to tavani.
reaches/she easily the ceiling
Giannis knows that Eleni said that that having long arms she
can easily reach the ceiling

(52) ?O Giannis kseri oti i Eleni ipe oti ftanei efkola
The G. knows that the/fem E. said that reaches/she easily
to tavani, echondas makria heria.
the ceiling having long arms
Giannis knows that Eleni said that he/she reaches the ceiling
easily, having long arm.

Stump (1985) points out that a subclass (his "*Weak*" Adjuncts) of modal gerunds generally behave like *if*-clauses.[19] In the above examples these correspond to the sentences in (44)-(47). We are interested here in their temporal interpretation and whether the patterns observed above hold also of this type of gerund clauses. This is indeed the case in (44)-(47) the adjunct can be construed with each one of the clauses in the complex structure. From this point of view then we can consider them as *when*-clauses, containing an empty temporal operator. This is not the case however in the examples (48)-(52) (Stump's "*strong*" Adjuncts). In these cases the adjunct can only be construed with the lowest clause. This difference can be traced to the stage/individual level status of the predicate. From the perspective of temporal interpretation, this fact does not undermine our proposal that there is a temporal operator, since, as I pointed out earlier, we have to account for the dependent temporal status of the adjunct. Stage-level predicates seem to allow the operator all possible scope options whereas individual-level predicates only admit narrowest scope. Consider however the effect of preposing the adjunct in (52) as in (53):

---

[19] Stump's discussion is broader. He considers all sorts of free adjuncts, including gerunds, we restrict here our attention on adjuncts of the latter type and consequently adapt some of his observations. We must also point out that Stump does not use our Manner - Temporal - Modal distinction which is intended to make more apparent the import of the syntax, provided that each part of the distinction corresponds to a specific syntactic configuration. Stump's aim rather is to discuss the interpretation of the apparently homogeneous class of *free adjuncts* from the points of view of Modality, Tense, and Aspect.

(53) Echondas makria heria, o Giannis kseri oti i Eleni ipe oti
having long arms the G. knows that the/fem E. said that
ftani efkola to tavani.
reaches/she easily the ceiling

In the most natural interpretation of (53) the adjunct is constructed with the matrix clause.[20] Consequently, in this case the operator must have wide scope. It seems that individual-level gerundival adjuncts have to be construed with the closest clause (downwards) rather than with the most deeply embedded as it would have been required if it had to take narrow scope. Somehow then this adjunct *belongs* to this clause in a more tight way. Why this is so? I want to propose here that in these cases the gerund is topicalised within its clause. It is moved to a Top position located at the complement of C. As it is natural, from this position the temporal operator, if this type of gerunds contain one, cannot move to the superordinate clause without violating the ECP. This proposal naturally explains some of the effects of the postposition of the adjunct as in (50). Assuming that the Top position is normally to the left of IP as shown also in Tsimpli (1992), (48) is ruled out as ungrammatical by the fact that the adjunct fails to be topicalised.[21] The question that this analysis raises is why only this type of gerund-adjuncts (strong adjuncts) must undergo topicalisation. Unfortunately I don't have a satisfactory answer to this question for the moment. Tentatively, I would like to suggest, as a first approximation, that the reason for this might have something to do with the fact that they derive from individual-level predicates whose interpretation is independent from any time intervals. They are somehow presupposed as topics generally are. Further refinements to this proposal are, no doubt, necessary. Space limitations prevent me from discussing this proposal further and I leave it for future research.

---

[20] It should be noted that (51) is judged somewhat strange by some speakers (including myself). I think this relative deviance is accountable on the nature of the predicate of each of the two clauses. The matrix predicate is stage level whereas the predicate of the embedded clause is individual level. Due partly to the embedded tense (habitual present) the embedded clause is interpreted as a generic sentence. Consequently, the modal gerund is more 'naturally' associated with the embedded rather than with the matrix, contrary to what is required by its position.

[21] Whether topicalisation involves movement or not is a question I will not address here. I will follow Chomsky 1977, Cinque 1991, Tsimpli 1992 in assuming that topicalised phrases are base-generated to their surface position, contrary to focused elements. My analysis would also be compatible with a movement approach to topicalisation if one wants to

To sum up, the syntactic behaviour of Modern Greek gerunds does not exactly parallel their semantic properties. They do not divide, syntactically into manner, temporal, and modal. Manner and temporal gerunds pattern in the same way as far as temporal interpretation is concerned and are opposed to modal gerunds.[22] The former show a considerable liberty in their temporal interpretation, which we accounted for by means of an abstract operator, whereas the latter are much more restricted in their scope options. The reason for this, I argued, is that they are topicalised in their clause.

4. Control in Gerunds

4.1 ECM, Argumenthood and the Subject of Gerunds

In this section I want to examine some issues arising with respect to the determination of the reference of the subject of gerund clauses in Modern Greek. Lexical subjects are generally not licensed in Modern Greek gerunds. As we saw above, gerund clauses can apparently never function as arguments. Therefore, it would be natural to suppose that they are never subject to Exceptional Case Marking. Therefore, even sentences like (54), which appear, *prima facie*, to be ECM structures

---

argue that argument topicalisation is different from adjunct topicalisation, for reasons such as predication

[22] Roughly speaking, this corresponds to Stump's Strong - Weak distinction.

have in fact to be distinct in some way or other from true ECM constructions.

(54) Thimamai ton Kosta odigondas to aftokinito.
Remember/I the K. driving the car
I remember Kostas driving the car.

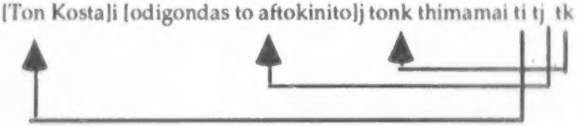
The DP *ton Kosta* can be cliticised on the main verb:

(55) Ton Thimamai odigondas to Aftokinito.
Him Remember/I driving the car
I remember him driving the car.

Furthermore, if the entire gerund, with the object, is topicalised then the object must obligatorily be linked to a resumptive preverbal clitic on the main verb ((56) and its schematic representation in (57)).[23] We can postulate that the clitic has moved to the preverbal position from its basic post-verbal position. This must be so since the only context in Modern Greek in which postverbal clitics are found is imperatives.

(56) Ton Kosta odigondas to aftokinito ton thimamai.
The K. driving the car HIM remember/I

(57)


[Ton Kosta]$_i$ [odigondas to aftokinito]$_j$ ton$_k$ thimamai t$_i$ t$_j$ t$_k$

Ton in (56) and (57) is the resumptive pronoun that the topicalised element is linked to. These can be considered as clitic doubling constructions.

---

[23] This is the standard pattern of Topicalisation in Modern Greek. See also Tsimpli 1992.

There are however some more difficult cases which tend to suggest that the DP object may in fact also be part of the gerund clause. Consider first, imperatives:

(58) a Ton Kosta odigondas to aftokinito thimisou.
The/Acc K driving the car remember/imp

b Ton Kosta odigondas to aftokinito thimisou ton.
The/Acc K. driving the car remember/imp him

c Ton Kosta odigondas to aftokinito thimisou to.
The/Acc K driving the car remember/imp it

d Ton Kosta thimisou ton odigondas to aftokinito.
The/Acc K. remember/imp him driving the car

Imperatives, which are the only context where the resumptive clitic could appear post-verbally, in fact show a different behaviour. In (58a) it is clear that what has been topicalised is one constituent, namely, the gerund clause. (58b) is what the sentence would have been had the only topicalised constituent been the object. Finally (58c) shows that the only way to express (58a) and still have a resumptive postverbal clitic would require the latter to be in the neuter form *to* 'it', corresponding to the meaning in (58e).

(58) e Remember the event (situation) in which Kostas was driving the car.

(58d) shows topicalisation of the object alone leaving the entire gerund clause behind. The following examples raise also the same problem:

(59) Ton Kosta odigondas to aftokinito (ton) ida ke trelathika.
The/Acc K driving the car (him) saw/I and went/I mad
I saw Kostas driving the car and went mad.

(60) Ton Kosta magirevondas (ton) thimithika ke eskasa sta gelia.
The/Acc K. cooking (him) remembered/I and burst/I in laughs
I remembered Kostas cooking and laughed.

These sentences show that, at least in some sense, our initial assumption, which is also the widely accepted view, that gerunds are always adjuncts and not subject to ECM is not accurate and must be revised in order to account for this restricted argument status of gerund clauses. It is restricted in the sense that only in some contexts, namely as complements to verbs selecting indefinite clausal complements, can they act as arguments.[24] The account of ECM that I am adopting here is the one presented in Tsoulas (forthcoming), and briefly outlined below: I take ECM to involve raising of the subject of the non-finite, Indefinite clausal complement to the specifier of the higher AgrO where it can check accusative Case. In order for this movement to be possible we must ensure that the Minimal Domain which this DP belongs to is properly extended. On the other hand I consider the selection of an Indefinite clausal complement as a marked selectional option,[25] therefore this feature (a head selects for a feature in the head of its complement) must be checked off. Checking the [+Indefinite] feature of the C head requires it to raise and adjoin to the selecting head, in a way similar to that in which Verb raises to T. It follows that the relevant Minimal Domain is extended accordingly, thus permitting the lower subject to raise to the specifier of AgrO.[26]

---

[24] It is precisely in those contexts in which they can alternate with subjunctives - the other type of indefinite clause one can find in Modern Greek. This is not true however cross-linguistically. It is not, for example, generally true for English. I have no explanation for this difference for the moment but I think it has to do with the fact that instead of infinitives Modern Greek possess only subjunctives, contrary to English. But I will not pursue this question any further here.

[25] I am considering any functional feature that has to be explicitly stated in the lexical entry of an item as a marked one.

[26] See Tsoulas 1995 for further technical details of this analysis.

Of course, in the vast majority of cases, when no lexical subject can be licensed in the adjunct the subject of the gerund is PRO.[27]

4.2 The Influence of the Object

I want now to turn back to the contrast mentioned in the introduction and consider the shifting in the control pattern in the light of the above discussion, consider again examples (1) and (2) repeated here:

(1) I Maria$_i$ ide to Gianni$_j$ [$_{CP}$ PRO$_{*i/j}$ zografizondas ena dendro].
The Maria saw the Gianni painting a tree
Maria saw Gianni while he was painting a tree.

(2) I Maria$_i$ ide to Gianni$_j$ [$_{CP}$ PRO$_{i/*j}$ zografizondas to dendro].
The Maria saw the Gianni painting the tree
Maria saw Gianni while she was painting the tree.

Given the above discussion it is natural to explain the quite puzzling contrast between (1) and (2) in terms of ECM, that is in (1) the verb *ide* Exceptionally Case marks inside the gerund clause, whereas this is somehow impossible in (2). I will argue that it is the presence of a definite object in (2) that is responsible for this situation. Recall that ECM depends on the indefinite nature of the clausal constituent. If the constituent is definite it is an absolute barrier to government and consequently ECM is precluded.[28] Thus, my proposal consists in the claim that the definiteness of the object is transferred to the gerund and furthermore to the entire CP. Krifka (1992) proposes a similar analysis of the *trade off* of grammatical features between verbal and nominal predicates affecting the temporal constitution of the sentence. As we saw at the beginning of this paper, gerunds differ from participles in several respects. We then considered participles as definites. Notice also that there are no active participles, morphologically distinguished as such, in Modern Greek. Transfer of a [+DEFINITE] feature to the gerund can be said to transform it into a more participle-like form, though somehow defective. This proposal, although very tentative and in need of considerable refinement, seems however quite accurate in that it also reflects the diachronic derivation of the gerund, which has presumably resulted in a form of ambiguity in the specifications of the *-ondas* morpheme.

---

[27] Although the presence, or absence, of PRO from the inventory of Modern Greek's grammatical categories is a rather controversial matter, no one, to the best of my knowledge, has ever suggested that PRO could be dispensed with in these constructions.

[28] Put in Minimalist terms, raising of the embedded subject to the superordinate Spec AgrO for accusative Case checking is impossible.

One possible objection to this analysis could be that apparently conflicting predictions are made by it and our analysis of the temporal interpretation of gerunds in terms of movement of an abstract operator. In fact the predictions are not conflicting because in one of those cases the gerund clause is an argument whereas in the other it is an adjunct. Of course, the question that still remains open is what happens with participles that are themselves adjuncts; also, why is it that only subject control is available in (2)? The answer to the latter question lies within the general mechanisms of Control theory. I would like to adopt here Williams' (1992) suggestion that in several cases of adjunct control, the controller is identified as the *logophoric centre of the sentence* in the case of (2) the perceiver is more likely to be the logophoric centre of the sentence in the sense of Sells (1987), and consequently the controller.

4. Conclusion

In this paper I have examined, as space limitations permitted, the structure and functioning of Modern Greek Gerundival constructions. I first argued that there are clear differences between gerunds and participles. I considered then issues concerning the temporal interpretation of gerunds and gave an account of it postulating the existence of a covert temporal operator akin to the one used by Geis (1970) for temporal prepositions in English, movement of this operator determines the clause with which the gerund will be associated. I assumed Kayne's (1994) theory of adjunction, which does not distinguish configurationally between adjunct phrases and specifiers in order to void a potential violation of the adjunct constraint (ECP). This analysis, independently, constitutes evidence for a disjunctive formulation of the ECP. I then considered issues of Control with gerunds and concluded that although apparently restricted to adjoined positions, gerunds can also be arguments and by virtue of their indefinite nature, they permit ECM. This partly resolves the problem raised by the sentences (1) and (2). On the other hand, following Krifka (1992) I argued that there is some feature transfer from the object to the gerund, which turns it to a more definite, participle-like constituent (but see note 25) which accounts for its control properties. The analysis presented in this paper represents further evidence for the Definite/Indefinite distinction at the clausal level. It should be noted however that the rather intuitive account of the properties of temporal/clausal indefiniteness given in this paper fails to do full justice to the linguistic reality it is supposed to account for.[29] In fact, temporal indefiniteness turns out to be much more complex than this intuitive account suggests. It also raises nontrivial questions, left untouched in this paper, concerning the representation of indefiniteness temporal or otherwise. Crucially, it sheds doubt on the widely accepted DRT idea of *Indefinites as variables* and it is possible that a detailed account of temporal indefiniteness will lead us to abandon this idea.[30] Additional reasons for such a move, from a Situation Semantics point of view, can be found in Cooper and Kamp (1991).

There are of course several other questions left open as indicated in the course of the paper. I leave all these questions for further research.

---

[29] See my 1994a, b, and forthcoming for some further details.

[30] However, Manzini 1994 presents ideas very similar to the ones presented in this paper and in Tsoulas 1994a, b and her analysis is fully cast in the framework of Heim's 1982 analysis of Indefinites-as-variables.

REFERENCES

Chierchia, G. 1984. *Topics in the Syntax and Semantics of Infinitives and Gerunds*. PhD Diss. UMass. Amherst.

Chomsky, N. 1977a. On WH-Movement. In Culicover, Peter et al. (eds.) *Formal Syntax*. New York, Academic Press.

Chomsky, N. 1977b. *Essays on Form and Interpretation*, Elsevier, North Holland.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht, Foris.

Chomsky, N. 1986a. *Barriers*, Cambridge, Mass. MIT Press.

Chomsky, N. 1986b. *Knowledge of Language*. New York, Praeger.

Chomsky, N. 1991. Some Notes on the Economy of Derivation and Representation. In Freidin, R. (ed.) *Principles and Parameters in Comparative Grammar*, Cambridge, Mass. MIT Press.

Chomsky, N. 1993. *A Minimalist Program for Linguistic Theory*. In Hale, K. and S. J. Keyser (eds.) *The view from Building 20*, Cambridge, Mass. MIT Press.

Chomsky, N. 1994. *Bare Phrase Structure*. MITOPL 4. Cambridge, MIT

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass. MIT Press

Cinque, G 1991: *Types of A'-dependencies*. Cambridge, Mass. MIT Press

Cooper, R and H Kamp. 1991. Negation in Situation Semantics and Discourse Representation Theory. In J Barwise et al. (eds.) *Situation Theory and its Applications vol.2* Stanford, CSLI.

Diesing, M. 1992. *Indefinites*. Cambridge, Mass. MIT Press.

Geis M. 1970. *Adverbial Subordinate Clauses in English*, PhD Diss. Cambridge, MIT.

Johnson, K. 1988. Clausal Gerunds, the ECP, and Government. *Linguistic Inquiry* 19.583-609.

Joseph, B. D. and I. Philippaki-Warburton 1986. *Modern Greek*, Croom Helm Descriptive Grammars, London, Croom Helm.

Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*, PhD Diss. UMass. Amherst.

Higginbotham, J. 1985. On Semantics. *Linguistic Inquiry* 16.547-593.

Higginbotham, J. 1992. Reference and Control. In Larson, R. K. et al. (eds.) *Control and Grammar*. Dordrecht, Kluwer. 79-108.

Householder, F. W., K. Kazazis and A. Koutsoudas 1964. *Reference Grammar of Literary Dhimotiki*. International Journal of American Linguistics. Publication 31 of the Indiana University Research Center in Anthropology, Folklore and Linguistics.

Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass. MIT Press.

Koopman, H. and Sportiche D. 1984. *Le Principe de Bijection*. In *Communications, 40* Paris, Editions du Seuil.

Krifka, M. 1992. Thematic Relations as Links between Nominal Reference and Temporal Constitution. In Sag, Ivan A. and Anna Szabolsci (eds.) *Lexical Matters*. Stanford, CSLI.

Laka, I. 1990. *Negation in Syntax: On the nature of functional categories and projections*. PhD Diss. Cambridge, MIT.

Manzini, M-R. 1993. *Locality: A theory and some of its empirical consequences*. Cambridge, Mass. MIT Press.

Manzini, M-R. 1994. *The Subjunctive*. Forthcoming in Nash L. and G. Tsoulas (eds.) *Paris-8 Working Papers in Linguistics* no. 1.

Mirambel, A. 1939. *Grammaire du Grec Moderne*. Paris Klincksieck.

Ross. J.-R. 1967. *Constraints on Variable in Syntax*. PhD Diss. Cambridge, MIT.

Portner, P. 1991. Interpreting Gerunds in Complement Positions, *Proceedings of the X^{th} West Coast Conference in Formal Linguistics* Stanford, CSLI.

Seiler, H. 1952. *L'aspect et le Temps dans le verbs Néo-Grec*. Paris, Les Belles Lettres.

Stowell, T. A. 1982. The Tense of Infinitives. *Linguistic Inquiry* 13.561-570.

Stowell, T. A. 1993. Syntax of Tense. Ms UCLA.

Stump, G. T. 1985. *The Semantic Variability of Absolute Constructions*. Dordrecht, Reidel.

Szabolcsi, A. 1989. Noun Phrases and Clauses: Is DP analogous to CP? Ms UCLA.

Terzi, A. 1992. *PRO in Finite Clauses. A study of the Inflectional Heads of the Balkan languages*. PhD Diss, CUNY.

Tsimpli, I. 1992. Focusing in Modern Greek, Ms, University College London.

Tsoulas, G. 1994a. Indefinite Clauses. Forthcoming in the Proceedings of the *XIIIth West Coast Conference on Formal Linguistics*, Stanford, CSLI.

Tsoulas, G. 1994b. Subjunctives as Indefinites. To appear in the proceedings of the *XX incontro de Grammatica Generativa*, Padova, Italy.

Tsoulas, G. 1994c. Minimalism and Control. Ms Université Paris-8.

Tsoulas, G. (forthcoming) The Nature of the Subjunctive and the Formal Grammar of Obviation. In K. Zagona (ed.) *Selected Papers from the XXVth Linguistic Symposium on Romance Languages*. Amsterdam, John Benjamins.

Tzartzanos, A. 1946, *Neoelliniki Sintaksis (ModernGreek Syntax)* Thessaloniki, Kiriakidis.

Williams, E. 1992. Adjunct Control. In Larson, R. K. et al. (eds.) *Control and Grammar*. Dordrecht, Kluwer. 297-322.

EDITORIAL STATEMENT

The editors welcome contributions for York Papers in Linguistics on any linguistic topic. Camera-ready copy is produced using Microsoft Word on an Apple Macintosh and it would be of considerable assistance to the editors if contributors could submit copy on a disk in this format. Failing this, Microsoft Word files on MS DOS format disks or MacWrite files on a Macintosh disk, together with a hard copy of the article, or an ASCII version of the text (in *any* disk format), again with accompanying hard copy, would be acceptable.

SJ Harlow
AR Warner
Series editors

STYLE SHEET FOR YORK PAPERS IN LINGUISTICS

Heading in capitals centralized and with * footnote for author's address, acknowledgements, etc. Then two blank lines and author centralized. Then one blank line and major affiliation. (No punctuation at ends of these lines). Then 2 blank lines and text.
Paragraphs indented. Use tabs for indent. No blank line between paragraphs. *Italics* for citation forms, single quotes for senses and for quotations, double quotes only for quotes within quotes.
Examples numbered consecutively. Numbers on the margin, in parentheses, with a., b., etc. for sub examples. Please use tabs before a., b., etc. and at the beginning (but only at the beginning) of example text and gloss, as follows

(1) a. J'ai lu [$_{NP}$ beaucoup d' articles] récemment
I've read many (of) articles recently

b. Pierre s'est brouillé avec [$_{NP}$ trop de collègues]
Pierre has argued with too-many (of) colleagues

Footnotes (after *) numbered consecutively and all placed at end of text.
References in the text in this form: Dowty (1982: 28). But omit parentheses in footnotes, and within other parentheses (like this: Dowty 1982: 28).

Major subheadings numbered and bold with word initial caps: number - stop - heading - no final punctuation, thus:
2. Fraggle Rock
One blank line before a major subheading.
Bibliography following footnotes, with centered-heading REFERENCES, and according to the conventions used in *Language* except that the date should be in parentheses, and the main title should be in italics and have initial capitals. Subordinate titles should be capitalized like ordinary text. Single quotes should *not* be used round titles:

Name, I. Q. (1998) *My Wonderful Book*. Someplace: Joe Publisher
Name, I. Q. (1998) A wonderful paper by me. *York Papers in Linguistics* 43.71-78.
Name, I. Q. (1998) A very wonderful paper by me. In Little, K. and Scrubb, J. (eds.) *Delights of Language*. Hopetown: Screwsome Press Inc. 324-328.

BACK ISSUES OF YORK PAPERS IN LINGUISTICS

The following are still available, though in some cases only two or three copies are left.

1992	YPL 16 (202pp.)	£8.
1991	YPL 15 (284pp.) (in honour of Jack Carnochan)	£10.
1989	YPL 14 (307pp.)	£10.
1989	YPL 13 (365pp.) (Festschrift for R B Le Page)	£10.
1986	YPL 12 (189pp.)	£5.
1984	YPL 11 (333pp.)	£5.
1983	YPL 10 (213pp.)	£5.
1976	YPL 6 (211pp.); 1975 YPL 5 (225pp.); 1974 YPL 4 (260pp.);	
1973	YPL 3 (183pp.)	£3. each

To order:
Send - details of the volume(s) you want
a cheque or draft in ***pounds sterling*** made payable to *University of York*
- your name and address

To - York Papers in Linguistics
Department of Language & Linguistic Science
University of York, Heslington
York Y01 5DD, U.K.

Prices include surface mail. Please add £3.50 per volume for air mail.